



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>6</sup> : <b>C12N 9/28, 15/56</b>	<b>A1</b>	(11) International Publication Number: <b>WO 96/23874</b> (43) International Publication Date: <b>8 August 1996 (08.08.96)</b>								
<p>(21) International Application Number: <b>PCT/DK96/00057</b></p> <p>(23) International Filing Date: <b>5 February 1996 (05.02.96)</b></p> <p>(30) Priority Data:</p> <table border="0"> <tr> <td>0128/95</td> <td>3 February 1995 (03.02.95)</td> <td>DK</td> </tr> <tr> <td>1192/95</td> <td>23 October 1995 (23.10.95)</td> <td>DK</td> </tr> <tr> <td>1256/95</td> <td>10 November 1995 (10.11.95)</td> <td>DK</td> </tr> </table> <p>(71) Applicant (for all designated States except US): <b>NOVO NORDISK A/S [DK/DK]; Novo Allé, DK-2880 Bagsvaerd (DK).</b></p> <p>(72) Inventors; and</p> <p>(73) Inventors/Applicants (for US only): <b>SVENDSEN, Allan [DK/DK]; Novo Nordisk a/s, Novo Allé, DK-2880 Bagsvaerd (DK). BISGÅRD-FRANTZEN, Henrik [DK/DK]; Novo Nordisk a/s, Novo Allé, DK-2880 Bagsvaerd (DK). BORCHERT, Torben, Vedel [DK/DK]; Novo Nordisk a/s, Novo Allé, DK-2880 Bagsvaerd (DK).</b></p> <p>(74) Common Representative: <b>NOVO NORDISK A/S; Corporate Patents, Novo Allé, DK-2880 Bagsvaerd (DK).</b></p>	0128/95	3 February 1995 (03.02.95)	DK	1192/95	23 October 1995 (23.10.95)	DK	1256/95	10 November 1995 (10.11.95)	DK	<p>(81) Designated States: <b>AL, AM, AT, AU, AZ, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AZ, BY, KG, KZ, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).</b></p> <p><b>Published</b></p> <p><i>With international search report.</i></p> <p><i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i></p>
0128/95	3 February 1995 (03.02.95)	DK								
1192/95	23 October 1995 (23.10.95)	DK								
1256/95	10 November 1995 (10.11.95)	DK								
(54) Title: <b>A METHOD OF DESIGNING ALPHA-AMYLASE MUTANTS WITH PREDETERMINED PROPERTIES</b>										
<p>(57) Abstract</p> <p>A method of constructing a variant of a parent Termamyl-like <math>\alpha</math>-amylase, which variant has <math>\alpha</math>-amylase activity and at least one altered property as compared to the parent <math>\alpha</math>-amylase, comprises i) analysing the structure of the parent Termamyl-like <math>\alpha</math>-amylase to identify at least one amino acid residue or at least one structural part of the Termamyl-like <math>\alpha</math>-amylase structure, which amino acid residue or structural part is believed to be of relevance for altering the property of the parent Termamyl-like <math>\alpha</math>-amylase (as evaluated on the basis of structural or functional considerations), ii) constructing a Termamyl-like <math>\alpha</math>-amylase variant, which as compared to the parent Termamyl-like <math>\alpha</math>-amylase, has been modified in the amino acid residue or structural part identified in i) so as to alter the property, and iii) testing the resulting Termamyl-like <math>\alpha</math>-amylase variant for the property in question.</p>										

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgyzstan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic	RR	Republic of Korea	SE	Sweden
CG	Congo	KZ	Kazakhstan	SG	Singapore
CH	Switzerland	LI	Liechtenstein	SI	Slovenia
CI	Côte d'Ivoire	LK	Sri Lanka	SK	Slovakia
CM	Cameroun	LR	Liberia	SN	Senegal
CN	China	LT	Lithuania	SS	Swaziland
CS	Czechoslovakia	LU	Luxembourg	TD	Chad
CZ	Czech Republic	LV	Latvia	TG	Togo
DE	Germany	MC	Monaco	TJ	Tajikistan
DK	Denmark	MD	Republic of Moldova	TT	Trinidad and Tobago
EE	Estonia	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	UG	Uganda
FI	Finland	MN	Mongolia	US	United States of America
FR	France	MR	Mauritania	UZ	Uzbekistan
GA	Gabon			VN	Viet Nam

## A METHOD OF DESIGNING ALPHA-AMYLASE MUTANTS WITH PREDETERMINED PROPERTIES

## FIELD OF THE INVENTION

5 The present invention relates to a novel method of designing  $\alpha$ -amylase mutants with predetermined properties, which method is based on the hitherto unknown three-dimensional structure of bacterial  $\alpha$ -amylases.

## 10 BACKGROUND OF THE INVENTION

$\alpha$ -Amylases ( $\alpha$ -1,4 glucan-4-glucanohydrolase, EC 3.2.1.1) constitute a group of enzymes which is capable of hydrolyzing starch and other linear and branched 1,4-glucosidic oligo- and  
15 polysaccharides. Almost all  $\alpha$ -amylases studied have a few conserved regions with approximately the same length and spacing. One of these regions resembles the  $\text{Ca}^{2+}$  binding site of calmodulin and the others are thought to be necessary for the active centre and/or binding of the substrate.

20

While the amino acid sequence and thus primary structure of a large number of  $\alpha$ -amylases are known, it has proved very difficult to determine the three-dimensional structure of all  $\alpha$ -amylases. The three-dimensional structure can be determined  
25 by X-ray crystallographic analysis of  $\alpha$ -amylase crystals, but it has proven difficult to obtain  $\alpha$ -amylase crystals suitable for actually solving the structure.

Until now the three-dimensional structure of only a few  
30  $\alpha$ -amylases have been determined at high resolution. These include the structure of the *Aspergillus oryzae* TAKA  $\alpha$ -amylase (Swift et al., 1991), the *Aspergillus niger* acid amylase (Brady et al., 1991), the structure of pig pancreatic  $\alpha$ -amylase (Qian et al., 1993), and the barley alpha-amylase (Kadziola et al.  
35 1994, Journal of Molecular Biology 239: 104-121, A.Kadziola, Thesis, Dept of Chemistry, U. of Copenhagen, Denmark). Furthermore, the three-dimensional structure of a *Bacillus circulans* cyclodextrin glycosyltransferase (CGTase) is known

(Klein et al., 1992) (Lawson et al., 1994). The CGTase catalyzes the same type of reactions as  $\alpha$ -amylases and exhibits some structural resemblance with  $\alpha$ -amylases.

5 Furthermore, crystallization and preliminary X-ray studies of *B. subtilis*  $\alpha$ -amylases have been described (Chang et al. (1992) and Mizuno et al. (1993)). No final *B. subtilis* structure has been reported. Analogously, the preparation of *B. licheniformis*  $\alpha$ -amylase crystals has been reported (Suzuki et al. (1990)), but  
10 no subsequent report on X-ray crystallographic analysis or three-dimensional structure are available.

Several research teams have attempted to build three-dimensional structures on the basis of the above known  
15  $\alpha$ -amylase structures. For instance, Vihinen et al. (J. Biochem. 107, 267-272, 1990), disclose the modelling (or computer simulation) of a three-dimensional structure of the *Bacillus stearothermophilus*  $\alpha$ -amylase on the basis of the TAKA amylase structure. The model was used to investigate hypothetical  
20 structural consequences of various site-directed mutations of the *B. stearothermophilus*  $\alpha$ -amylase. E.A. MacGregor (1987) predicts the presence of  $\alpha$ -helices and  $\beta$ -barrels in  $\alpha$ -amylases from different sources, including barley, pig pancreas and *Bacillus amyloliquefaciens* on the basis of the known structure  
25 of the *A. oryzae* TAKA  $\alpha$ -amylase and secondary structure predicting algorithms. Furthermore, the possible loops and subsites which may be found to be present in, e.g., the *B. amyloliquefaciens*  $\alpha$ -amylase are predicted (based on a comparison with the *A. oryzae* sequence and structure).

30

A.E. MacGregor (Starch/Stärke 45 (1993), No. 7, p. 232-237) presents a review of the relationship between the structure and activity of  $\alpha$ -amylase related enzymes.

35 Hitherto, no three-dimensional structure has been available for the industrially important *Bacillus*  $\alpha$ -amylases (which in the present context are termed "Termamyl-like  $\alpha$ -amylases"),



including the *B. licheniformis*, the *B. amyloliquefaciens*, and the *B. stearothermophilus*  $\alpha$ -amylase.

#### BRIEF DISCLOSURE OF THE INVENTION

5

The three-dimensional structure of a Termamyl-like bacterial  $\alpha$ -amylase has now been elucidated. On the basis of an analysis of said structure it is possible to identify structural parts or specific amino acid residues which from structural or  
10 functional considerations appear to be important for conferring the various properties to the Termamyl-like  $\alpha$ -amylases. Furthermore, when comparing the Termamyl-like  $\alpha$ -amylase structure with known structures of the fungal and mammalian  $\alpha$ -amylases mentioned above, it has been found that some  
15 similarities exist between the structures, but also that some striking, and not previously predicted structural differences between the  $\alpha$ -amylases exist. The present invention is based on these findings.

20 Accordingly, in a first aspect the invention relates to a method of constructing a variant of a parent Termamyl-like  $\alpha$ -amylase, which variant has  $\alpha$ -amylase activity and at least one altered property as compared to said parent  $\alpha$ -amylase, which method comprises

25

i) analysing the structure of the Termamyl-like  $\alpha$ -amylase with a view to identifying at least one amino acid residue or at least one structural part of the Termamyl-like  $\alpha$ -amylase structure, which amino acid residue or structural part is  
30 believed to be of relevance for altering said property of the parent Termamyl-like  $\alpha$ -amylase (as evaluated on the basis of structural or functional considerations),

ii) constructing a Termamyl-like  $\alpha$ -amylase variant, which as  
35 compared to the parent Termamyl-like  $\alpha$ -amylase, has been modified in the amino acid residue or structural part identified in i) so as to alter said property, and

iii) testing the resulting Termamyl-like  $\alpha$ -amylase variant for said property.

- 5 In a second aspect the present invention relates to a method of constructing a variant of a parent Termamyl-like  $\alpha$ -amylase, which variant has  $\alpha$ -amylase activity and one or more altered properties as compared to said parent  $\alpha$ -amylase, which method comprises
- 10 i) comparing the three-dimensional structure of the Termamyl-like  $\alpha$ -amylase with the structure of a non-Termamyl-like  $\alpha$ -amylase,
- ii) identifying a part of the Termamyl-like  $\alpha$ -amylase structure which is different from the non-Termamyl-like  $\alpha$ -amylase
- 15 structure, and
- iii) modifying the part of the Termamyl-like  $\alpha$ -amylase identified in ii) whereby a Termamyl-like  $\alpha$ -amylase variant is obtained, one or more properties of which differ from the parent Termamyl-like  $\alpha$ -amylase.

20

In a third aspect the invention relates to a method of constructing a variant of a parent non-Termamyl-like  $\alpha$ -amylase, which variant has  $\alpha$ -amylase activity and one or more altered properties as compared to said parent  $\alpha$ -amylase, which method

25 comprises

- i) comparing the three-dimensional structure of the non-Termamyl-like  $\alpha$ -amylase with the structure of a Termamyl-like  $\alpha$ -amylase,
- ii) identifying a part of the non-Termamyl-like  $\alpha$ -amylase
- 30 structure which is different from the Termamyl-like  $\alpha$ -amylase structure, and
- iii) modifying the part of the non-Termamyl-like  $\alpha$ -amylase identified in ii) whereby a non-Termamyl-like  $\alpha$ -amylase variant is obtained, one or more properties of which differ from the
- 35 parent Termamyl-like  $\alpha$ -amylase.

The property which may be altered by the above methods of the present invention may, e.g., be substrate specificity,

substrate binding, substrate cleavage pattern, temperature stability, pH dependent activity, pH dependent stability (especially increased stability at low (e.g. pH<6, in particular pH<5) or high (e.g. pH>9) pH values), stability towards oxidation, Ca<sup>2+</sup>-dependency, specific activity, and other properties of interest. For instance, the alteration may result in a variant which, as compared to the parent Termamyl-like  $\alpha$ -amylase, has an increased specific activity at a given pH and/or an altered substrate specificity.

10

In still further aspects the invention relates to variants of a Termamyl-like  $\alpha$ -amylase, DNA encoding such variants and methods of preparing the variants. Finally, the invention relates to the use of the variants for various industrial purposes.

15

#### DETAILED DISCLOSURE OF THE INVENTION

##### The Termamyl-like $\alpha$ -amylase

20

It is well known that a number of alpha-amylases produced by *Bacillus* spp. are highly homologous on the amino acid level. For instance, the *B. licheniformis*  $\alpha$ -amylase comprising the amino acid sequence shown in SEQ ID No. 2 (commercially available as Termamyl®) has been found to be about 89% homologous with the *B. amyloliquefaciens*  $\alpha$ -amylase comprising the amino acid sequence shown in SEQ ID No. 4 and about 79% homologous with the *B. stearothermophilus*  $\alpha$ -amylase comprising the amino acid sequence shown in SEQ ID No. 6. Further homologous  $\alpha$ -amylases include an  $\alpha$ -amylase derived from a strain of the *Bacillus* sp. NCIB 12289, NCIB 12512, NCIB 12513 or DSM 9375, all of which are described in detail in WO 95/26397, and the  $\alpha$ -amylase described by Tsukamoto et al., 1988, Biochemical and Biophysical Research Communications, Vol. 151, No. 1. Still other homologous  $\alpha$ -amylases include the  $\alpha$ -amylase produced by the *B. licheniformis* described in EP 252 666 (ATCC 27811), and the  $\alpha$ -amylases identified in WO 91/00353 and WO 94/18314. Other commercial Termamyl-like *B.*

30

35

*licheniformis*  $\alpha$ -amylases are Optitherm<sup>®</sup> and Takatherm<sup>®</sup> (available from Solvay), Maxamyl<sup>®</sup> (available from Gist-brocades/Genencor), Spezym AA<sup>®</sup> (available from Genencor), and Keistase<sup>®</sup> (available from Daiwa).

5

Because of the substantial homology found between these  $\alpha$ -amylases, they are considered to belong to the same class of  $\alpha$ -amylases, namely the class of "Termamyl-like  $\alpha$ -amylases".

- 10 Accordingly, in the present context, the term "Termamyl-like  $\alpha$ -amylase" is intended to indicate an  $\alpha$ -amylase which, on the amino acid level, exhibits a substantial homology to Termamyl<sup>®</sup>, i.e. the *B. licheniformis*  $\alpha$ -amylase SEQ ID NO 2. In other words, a Termamyl-like  $\alpha$ -amylase is an  $\alpha$ -amylase, which has the
- 15 amino acid sequence shown in SEQ ID No. 2, 4 or 6 herein, or the amino acid sequence shown in SEQ ID NO 1 or 2 of WO 95/26397 or in Tsukamoto et al., 1988, or i) which displays at least 60%, such as at least 70%, e.g. at least 75%, or at least 80%, e.g. at least 85%, at least 90% or at least 95% homology
- 20 with at least one of said amino acid sequences and/or ii) displays immunological cross-reactivity with an antibody raised against at least one of said  $\alpha$ -amylases, and/or iii) is encoded by a DNA sequence which hybridizes to the DNA sequences encoding the above specified  $\alpha$ -amylases which are apparent from
- 25 SEQ ID Nos. 1, 3 and 5 of the present application, and SEQ ID NO 4 and 5 of WO 95/26397, respectively.

In connection with property i) the "homology" may be determined by use of any conventional algorithm, preferably by use of the

30 GAP programme from the GCG package version 7.3 (June 1993) using default values for GAP penalties (Genetic Computer Group (1991) Programme Manual for the GCG Package, version 7, 575 Science Drive, Madison, Wisconsin, USA 53711).

- 35 Property ii) of the  $\alpha$ -amylase, i.e. the immunological cross reactivity, may be assayed using an antibody raised against or reactive with at least one epitope of the relevant Termamyl-like  $\alpha$ -amylase. The antibody, which may either be monoclonal or

polyclonal, may be produced by methods known in the art, e.g. as described by Hudson et al., 1989. The immunological cross-reactivity may be determined using assays known in the art, examples of which are Western Blotting or radial immunodiffusion assay, e.g. as described by Hudson et al., 1989. In this respect, immunological cross-reactivity between the  $\alpha$ -amylases having the amino acid sequences SEQ ID Nos. 2, 4 and 6, respectively, has been found.

- 10 The oligonucleotide probe used in the characterization of the Termamyl-like  $\alpha$ -amylase in accordance with property (iii) above may suitably be prepared on the basis of the full or partial nucleotide or amino acid sequence of the  $\alpha$ -amylase in question. Suitable conditions for testing hybridization involve  
15 presoaking in 5xSSC and prehybridizing for 1h at  $-40^{\circ}\text{C}$  in a solution of 20% formamide, 5xDenhardt's solution, 50mM sodium phosphate, pH 6.8, and 50 $\mu\text{g}$  of denatured sonicated calf thymus DNA, followed by hybridization in the same solution supplemented with 100 $\mu\text{M}$  ATP for 18h at  $-40^{\circ}\text{C}$ , or other methods  
20 described by e.g. Sambrook et al., 1989.

In the present context, "derived from" is intended not only to indicate an  $\alpha$ -amylase produced or producible by a strain of the organism in question, but also an  $\alpha$ -amylase encoded by a DNA  
25 sequence isolated from such strain and produced in a host organism transformed with said DNA sequence. Finally, the term is intended to indicate an  $\alpha$ -amylase which is encoded by a DNA sequence of synthetic and/or cDNA origin and which has the identifying characteristics of the  $\alpha$ -amylase in question. The  
30 term is also intended to indicate that the parent  $\alpha$ -amylase may be a variant of a naturally occurring  $\alpha$ -amylase, i.e. a variant which is the result of a modification (insertion, substitution, deletion) of one or more amino acid residues of the naturally occurring  $\alpha$ -amylase.

Parent hybrid  $\alpha$ -amylases

The parent  $\alpha$ -amylase (being a Termamyl-like or non-Termamyl-like  $\alpha$ -amylase) may be a hybrid  $\alpha$ -amylase, i.e. an  $\alpha$ -amylase which comprises a combination of partial amino acid sequences derived from at least two  $\alpha$ -amylases.

The parent hybrid  $\alpha$ -amylase may be one which on the basis of amino acid homology and/or immunological cross-reactivity and/or DNA hybridization (as defined above) can be determined to belong to the Termamyl-like  $\alpha$ -amylase family. In this case, the hybrid  $\alpha$ -amylase is typically composed of at least one part of a Termamyl-like  $\alpha$ -amylase and part(s) of one or more other  $\alpha$ -amylases selected from Termamyl-like  $\alpha$ -amylases or non-Termamyl-like  $\alpha$ -amylases of microbial (bacterial or fungal) and/or mammalian origin.

Thus, the parent hybrid  $\alpha$ -amylase may comprise a combination of at least two Termamyl-like  $\alpha$ -amylases, or of at least one Termamyl-like and at least one non-Termamyl-like bacterial  $\alpha$ -amylase, or of at least one Termamyl-like and at least one fungal  $\alpha$ -amylase. For instance, the parent  $\alpha$ -amylase comprises a C-terminal part of an  $\alpha$ -amylase derived from a strain of *B. licheniformis* and a N-terminal part of an  $\alpha$ -amylase derived from a strain of *B. amyloliquefaciens* or from a strain of *B. stearothermophilus*. For instance, the parent  $\alpha$ -amylase comprises at least 430 amino acid residues of the C-terminal part of the *B. licheniformis*  $\alpha$ -amylase, and may, e.g. comprise a) an amino acid segment corresponding to the 37 N-terminal amino acid residues of the *B. amyloliquefaciens*  $\alpha$ -amylase having the amino acid sequence shown in SEQ ID No. 4 and an amino acid segment corresponding to the 445 C-terminal amino acid residues of the *B. licheniformis*  $\alpha$ -amylase having the amino acid sequence shown in SEQ ID No. 2, or b) an amino acid segment corresponding to the 68 N-terminal amino acid residues of the *B. stearothermophilus*  $\alpha$ -amylase having the amino acid sequence shown in SEQ ID No. 6 and an amino acid segment corresponding to the 415 C-terminal amino acid residues of the

*B. licheniformis*  $\alpha$ -amylase having the amino acid sequence shown in SEQ ID No. 2.

Analogously, the parent hybrid  $\alpha$ -amylase may belong to a non-Termamyl-like  $\alpha$ -amylase family, e.g. the Fungamyl-like  $\alpha$ -amylase family. In that case the hybrid may comprise at least one part of an  $\alpha$ -amylase belonging to the non-Termamyl-like  $\alpha$ -amylase family in combination with one or more parts derived from other  $\alpha$ -amylases.

10

The three-dimensional Termamyl-like  $\alpha$ -amylase structure

The Termamyl-like  $\alpha$ -amylase which was used to elucidate the three-dimensional structure forming the basis for the present invention consists of the 300 N-terminal amino acids of the *B. amyloliquefaciens*  $\alpha$ -amylase (with the amino acid sequence shown in SEQ ID No. 4) and amino acids 301-483 of the C-terminal end of the *B. licheniformis*  $\alpha$ -amylase with the amino acid sequence SEQ ID No. 2. The bacterial  $\alpha$ -amylase belongs to the "Termamyl-like  $\alpha$ -amylase family" and the present structure is believed to be representative for the structure of any Termamyl-like  $\alpha$ -amylase.

The structure of the  $\alpha$ -amylase was solved in accordance with the principle for X-ray crystallographic methods given in "X-Ray Structure Determination", Stout, G.K. and Jensen, L.H., John Wiley & Sons, inc. NY, 1989. The structural coordinates for the solved crystal structure of the  $\alpha$ -amylase at 2.2 Å resolution using the isomorphous replacement method are given in a standard PDB format (Brookhaven Protein Data Base) in Appendix 1. It is to be understood that Appendix 1 forms part of the present application.

Amino acid residues of the enzyme are identified by three-letter amino acid code (capitalized letters).

The  $\alpha$ -amylase structure is made up of three globular domains ordered A, B, and C with respect to sequence, which lie

approximately along a line in the order B, A, C. The domains can be defined as being residues 1-103 and 206-395 for domain A, residues 104-205 for domain B, and residues 396-483 for domain C, the numbers referring to the *B. licheniformis*  $\alpha$ -amylase. This gives rise to an elongated molecule, the longest axis being about 85Å. The widest point perpendicular to this axis is approximately 50Å and spans the central A domain. The active site residues of the *B. licheniformis*  $\alpha$ -amylase (SEQ ID NO 2) are D323, D231 and E261.

10

#### Domain A

Domain A is the largest domain and contains the active site (comprised of a cluster of three amino acid residues placed at the bottom of a deep cleft in the enzyme's surface). Domain A of all known  $\alpha$ -amylase structures have the same overall fold, viz. the (beta/alpha)<sub>8</sub> barrel with 8 central beta strands (number 1-8) and 8 flanking  $\alpha$ -helices. The  $\beta$ -barrel is defined by McGregor *op. cit.* The C-terminal end of Beta strand 1 is connected to helix 1 by a loop denoted loop 1 and an identical pattern is found for the other loops. These loops show some variation in size and some can be quite extensive.

The 8 central Beta-strands in the (beta/alpha)<sub>8</sub> barrel superimpose well between the various known  $\alpha$ -amylase structures, and this part of the structure, including the close surroundings of the active site located at the c-terminal end of the beta-strands, show high similarity between the different amylases.

30

The loops connecting beta-strands and alpha helices display high variations between alpha amylases. These loops constitute the structural context of the active site and the majority of the contacts to the substrate is found among residues located in these loops. Such important characteristics as substrate specificity, substrate binding, pH/activity profile, starch cleavage pattern are determined by the amino acids and the positions of same in these loops.



The substantial differences between the Fungamyl-like  $\alpha$ -amylase structure and the structure of the Termamyl-like  $\alpha$ -amylase disclosed herein which are found in loops 1, 2, 3, and 8 are visualized in the Figures.

5

#### Domain B

The Termamyl-like  $\alpha$ -amylase structure has been found to comprise a special domain structure in the A domain's loop3, also called domain B. The structure of the Termamyl-like  $\alpha$ -amylase B domain has never been seen before in any of the known  $\alpha$ -amylase or  $(\beta/\alpha)_8$ -barrel proteins.

The domain B structure is a very compact domain having a very high number of charged residues. The B domain arises as an extension of the loop between strand 3 and helix 3 of domain A (shown in Fig. 7) and contains a 5 stranded antiparallel  $\beta$ -sheet structure containing at least one long loop structure and having the connectivity -1, +3, -1X, +2 (Richardson, 1981, Adv. Protein Chem. 34, 167-339).

The first four strands of the B domain form two hairpin loops which twist around each other like a pair of crossed fingers (right-hand twist). The mainchain folds into a  $\beta$ -strand which connects two small  $\beta$ -sheet structures. After making one turn in one sheet it folds back and makes up a two stranded sheet in contact with domain A and an internal hole in the  $\alpha$ -amylase structure. Then the mainchain folds up to a small sheet structure nearly perpendicular to the first two sheets. Before entering the helix 3 on top of the  $\beta$ -strand 3, the approximately 24 last amino acids in domain B form two calcium binding sites in the contact region to domain A.

Domain B is connected with domain A by two peptide stretches, which divide the domain-domain contact areas into two. Domain B is in contact with Domain A by a calcium binding region and an internally buried hole containing waters. Many types of molecular contacts are present. Ionic interacting between acid

and basic amino acids are possible, these interactions are very important for the general stability at high pH and for keeping the Calcium binding sites intact.

#### 5 Domain C

Domain C is the C-terminal part of the protein consisting of amino acids 394-483. Domain C is composed entirely of  $\beta$ -strands which forms a single 8-stranded sheet structure, which folds  
10 back on itself, and thus may be described as a  $\beta$ -sandwich structure. The connectivity is +1,+1, +5, -3, +1, +1, -3 although strands 6 and 7 are only loosely connected. One part of the  $\beta$ -sheet forms the interface to domain A.

#### 15 Ca-binding and Na-binding sites

The structure of the Termamyl-like  $\alpha$ -amylase is remarkable in that it exhibits four calcium-binding sites and one sodium-binding site. In other words four calcium ions and one sodium  
20 ion are found to be present in the structure, although one of the calcium ions displays very weak coordination. Two of the calcium ions form part of a linear cluster of three ions, the central ion being attributed to sodium, which lie at the junction of the A and B domains.

25

The coordinating residues for the calcium ions between the A and B domain are as follows (using the Pdb file nomenclature for amino acid residues and atoms in the Pdb file found in Appendix 1 herein): For the calcium ion nearest to the active  
30 site (IUM 502 in the pdb file), the backbone carbonyls from His235 and Asp194, the sidechain atom OD1 from residues Asp194, Asn102 and Asp200, and one water molecule WAT X3 (atom OW7). For the sodium ion (IUM 505), the binding site includes atom OD2 from Asp194, Asp200, Asp183 and Asp159, and a backbone  
35 carbonyl from Val201. The coordinates for the other calcium ion between domain A and B are (IUM 501) : atom OD2 from Asp204 and Asp159, backbone carbonyl from Asp183 and Ala181, atom OD1 from Asp202, and one water molecule WAT X7 (atom OW7).

One calcium ion is located between the A and C domain, another is located in the C domain. The first mentioned calcium ion, which is also the one best coordinated (IUM 503) includes a carbonyl backbone from Gly300, Tyr302 and His406, atom OD2/OD1 from Asp430, atom OD1 from Asp407, and one water molecule WAT X6 (atom OW7). The other and very weakly coordinated calcium site (IUM 504) comprises 4 water molecules WAT X21 (atom OW8), X6 (atom OW6), X9 (atom OW0) and X28 (atom OW8), OE1/OE2 from Glu447 and OD1 from Asn444.

10

#### *Substrate-binding site*

Without being limited to any theory it is presently believed that favourable interactions between a substrate molecule and the enzyme (such as hydrogen bonds and/or strong electrostatic interaction) are found within a sphere of 4Å of the substrate, when bound to the enzyme. The following residues of the *B. licheniformis*  $\alpha$ -amylase having the amino acid sequence shown in SEQ ID No. 2 are contemplated to be within a distance of 4 Å of the substrate and thus believed to be involved in interactions with the substrate:

Trp13, Tyr14, Asn17, Asp18, Ser50, Gln51, Ala52, Asp53, Val54, Gly55, Tyr56, Lys70, Arg74, Lys76, Val102, His105, Gly107, Gly108, Ala109, Trp138, Thr163, Asp164, Trp165, Asn172, Glu189, Tyr193, Leu196, Met197, Tyr198, Ala199, Arg229, Asp231, Ala232, Lys234, His235, Glu261, Trp263, His327, Asp328, Gln333, Ser334, and Leu335.

The amino acid residues of another Termamyl-like  $\alpha$ -amylase, which are contemplated to be within a distance of 4Å of the substrate, may easily be identified by aligning the amino acid sequence SEQ ID NO 2 with that of the other Termamyl-like  $\alpha$ -amylase and thereby identifying the positions equivalent to those identified above.

15

Generality of structure

Because of the high homology between the various Termamyl-like  $\alpha$ -amylases, the solved structure defined by the coordinates of Appendix 1 is believed to be representative for the structure of all Termamyl-like  $\alpha$ -amylases. A model structure of other Termamyl-like  $\alpha$ -amylases may easily be built on the basis of the coordinates given in Appendix 1 adapted to the  $\alpha$ -amylase in question by use of an alignment between the respective amino acid sequences. The creation of a model structure is exemplified in Example 1.

The above identified structurally characteristic parts of the Termamyl-like  $\alpha$ -amylase structure (Ca-binding site, substrate binding site, loops, etc.) may easily be identified in other Termamyl-like  $\alpha$ -amylases on the basis of a model (or solved) structure of the relevant Termamyl-like  $\alpha$ -amylase or simply on the basis of an alignment between the amino acid sequence of the Termamyl-like  $\alpha$ -amylase in question with that of the *B. lichenformis*  $\alpha$ -amylase used herein for identifying the amino acid residues of the respective structural elements.

Furthermore, in connection with Termamyl-like variants of the invention, which are defined by modification of specific amino acid residues of a specific Termamyl-like  $\alpha$ -amylase, it will be understood that variants of another Termamyl-like  $\alpha$ -amylase modified in an equivalent position (as determined from the best possible amino acid sequence alignment between the respective sequences) are intended to be covered as well. Thus, irrespective of whether an amino acid residue is identified herein for the purpose of defining a structural part of a given  $\alpha$ -amylase or used for identifying a variant of the  $\alpha$ -amylase, this amino acid residue shall be considered as representing the equivalent amino acid residue of any other Termamyl-like  $\alpha$ -amylase.

Methods of the invention for design of novel  $\alpha$ -amylase variants

In the methods according to the first, second and third aspects of the invention the terms "structure of a Termamyl-like  $\alpha$ -amylase" and "Termamyl-like  $\alpha$ -amylase structure" are intended to indicate the solved structure defined by the coordinates presented in Appendix 1 or a model structure of a given Termamyl-like  $\alpha$ -amylase (such as the *B. licheniformis*  $\alpha$ -amylase) built on the basis of the solved structure.

10

In most cases the parent Termamyl-like  $\alpha$ -amylase to be modified in accordance with the present invention is different from the  $\alpha$ -amylase which was actually used for solving the structure (Appendix 1). This means that the amino acid residue(s) or structural part(s) identified in the solved structure (Appendix 1) in step i) of the method according to the first, second or third aspect of the invention must be translated into the corresponding amino acid residue(s) or structural part(s) of the parent Termamyl-like  $\alpha$ -amylase in question. The "translation" is conveniently performed on the basis of an amino acid sequence alignment between the amino acid sequence of the Termamyl-like  $\alpha$ -amylase used for solving the structure and the amino acid sequence of the parent Termamyl-like  $\alpha$ -amylase in question.

25

The analysis or comparison performed in step i) of the method according to the first, second and third aspect, respectively, of the invention may be performed by use of any suitable computer programme capable of analysing and/or comparing protein structures, e.g. the computer programme Insight, available from Biosym Technologies, Inc. For instance, the basic principle of structure comparison is that the three-dimensional structures to be compared are superimposed on the basis of an alignment of secondary structure elements (such as the central 8  $\beta$ -strands in the barrel) and the parts differing between the structures can subsequently easily be identified from the superimposed structure.

The structural part which is identified in step i) of the methods of the first, second and third aspects of the invention may be composed of one amino acid residue. However, normally the structural part comprises more than one amino acid residue, typically constituting one of the above parts of the Termamyl-like  $\alpha$ -amylase structure such as one of the A, B, or C domains, an interface between any of these domains, a calcium binding site, a loop structure, the substrate binding site, or the like.

10

In the present context the term "structural or functional considerations" is intended to indicate that modifications are made on the basis of an analysis of the relevant structure or structural part and its contemplated impact on the function of the enzyme. Thus, an analysis of the structures of the various  $\alpha$ -amylases, which until now has been elucidated, optionally in combination with an analysis of the functional differences between these  $\alpha$ -amylases, may be used for assigning certain properties of the  $\alpha$ -amylases to certain parts of the  $\alpha$ -amylase structure or to contemplate such relationship. For instance, differences in the pattern or structure of loops surrounding the active site may result in differences in access to the active site of the substrate and thus differences in substrate specificity and/or cleavage pattern. Furthermore, parts of a Termamyl-like  $\alpha$ -amylase involved in or contemplated to be involved in substrate binding (and thus e.g. specificity/cleavage pattern), calcium or sodium ion binding (e.g. of importance for the Calcium-dependency of the enzyme), and the like has been identified (*vide infra*).

30

The modification of an amino acid residue or structural part is typically accomplished by suitable modifications of a DNA sequence encoding the parent enzyme in question. The term "modified" as used in step ii) in the method according to the first aspect of the invention is intended to have the following meaning: When used in relation to an amino acid residue the term is intended to mean replacement of the amino acid residue in question with another amino acid residue. When used in

relation to a structural part, the term is intended to mean replacement of one or more amino acid residues of said structural part, addition of one or more amino acid residues to said part, or deletion of one or more amino acid residues of  
5 said structural part.

The construction of the variant of interest is accomplished by cultivating a microorganism comprising a DNA sequence encoding the variant under conditions which are conducive for producing  
10 the variant, and optionally subsequently recovering the variant from the resulting culture broth. This is described in detail further below.

#### *First aspect of the invention*

15 In a preferred embodiment of the method according to the first aspect of the invention the property of the parent enzyme to be modified is selected from calcium dependency, substrate binding, cleavage pattern, pH dependent activity and the like. Specific examples of how to change these properties of a parent  
20 Termamyl-like  $\alpha$ -amylase are given further below.

In another preferred embodiment the parent Termamyl-like  $\alpha$ -amylase to be modified is a *B. licheniformis*  $\alpha$ -amylase.

#### *25 Second and third aspects of the invention*

One important advantage of the methods according to the second and third aspects of the present invention is that it is possible to adapt the structure (or a structural part) of a Termamyl-like  $\alpha$ -amylase to the structure (or structural part)  
30 of a non-Termamyl-like  $\alpha$ -amylase and vice versa. For instance, having identified a loop structure of the non-Termamyl-like  $\alpha$ -amylase which is believed to be responsible for or contributing to a particular property of the non-Termamyl-like  $\alpha$ -amylase it is possible to replace the corresponding structure of the  
35 Termamyl-like  $\alpha$ -amylase with said non-Termamyl-like  $\alpha$ -amylase structure - or if no corresponding structure exists in the Termamyl-like  $\alpha$ -amylase - to insert the structure into the Termamyl-like  $\alpha$ -amylase in such a manner that the resulting

variant Termamyl-like  $\alpha$ -amylase, as far as the relevant part is concerned, resembles the corresponding part of the non-Termamyl-like  $\alpha$ -amylase. When two or more parts of the structure of the parent Termamyl-like  $\alpha$ -amylase are modified so  
5 as to resemble the corresponding parts of the non-Termamyl-like  $\alpha$ -amylase it is possible to increase the resemblance to the non-Termamyl-like  $\alpha$ -amylase of the Termamyl-like  $\alpha$ -amylase variant and thus to alter the properties of said variant in the direction of those of said non-Termamyl-like  $\alpha$ -amylase. Loop  
10 modifications are discussed in much further detail further below.

Typically, the modification to be performed in step iii) of the method according to the second aspect of the invention is  
15 accomplished by deleting one or more amino acid residues of the part of the Termamyl-like  $\alpha$ -amylase to be modified so as to adapt the structure of said part of the parent  $\alpha$ -amylase to the corresponding part of the non-Termamyl-like  $\alpha$ -amylase; by replacing one or more amino acid residues of the part of the  
20 Termamyl-like  $\alpha$ -amylase to be modified with the amino acid residues occupying corresponding positions in the non-Termamyl-like  $\alpha$ -amylase; or by insertion of one or more amino acid residues present in the non-Termamyl-like  $\alpha$ -amylase into a corresponding position in the Termamyl-like  $\alpha$ -amylase. For the  
25 method according to the third aspect the modification is to be understood analogously, performed on the non-Termamyl-like parent  $\alpha$ -amylase rather than the Termamyl-like  $\alpha$ -amylase.

In step ii) of the method according to the second or third  
30 aspect of the invention the part of the structure to be identified is preferably one which in the folded enzyme is believed to be in contact with the substrate (cf the disclosure above in the section entitled "Substrate-binding site) or involved in substrate specificity and/or cleavage pattern,  
35 and/or one which is in contact with one of the calcium or sodium ions and/or one, which is contributing to the pH or temperature profile of the enzyme, or one which otherwise, from structural or functional considerations, is contemplated to be



responsible for differences in one or more properties of the Termamyl-like and non-Termamyl-like  $\alpha$ -amylase.

*Non-Termamyl-like  $\alpha$ -amylase*

5 The non-Termamyl-like  $\alpha$ -amylase with which the comparison is made in step i) of the method of the second aspect of the invention and which is the parent  $\alpha$ -amylase in the method of the third aspect of the invention, may be any  $\alpha$ -amylase, which does not belong to the family of Termamyl-like  $\alpha$ -amylases (as  
10 defined above) and, which as a consequence thereof, has a different three-dimensional structure. Furthermore, the non-Termamyl-like  $\alpha$ -amylase should be one which has, at the time that the method is performed, an elucidated or contemplated three-dimensional structure.

15 The non-Termamyl-like  $\alpha$ -amylase may, e.g., be a fungal  $\alpha$ -amylase, a mammalian or a plant  $\alpha$ -amylase or a bacterial  $\alpha$ -amylase (different from a Termamyl-like  $\alpha$ -amylase). Specific examples of such  $\alpha$ -amylases include the *Aspergillus oryzae* TAKA  
20  $\alpha$ -amylase, the *A. niger* acid  $\alpha$ -amylase, the *Bacillus subtilis*  $\alpha$ -amylase, the porcine pancreatic  $\alpha$ -amylase and a barley  $\alpha$ -amylase. All of these  $\alpha$ -amylases have elucidated structures which are clearly different from the structure of the Termamyl-like  $\alpha$ -amylase shown herein.

25 The fungal  $\alpha$ -amylases mentioned above, i.e. derived from *A. niger* and *A. oryzae*, are highly homologous on the amino acid level and generally considered to belong to the same family of  $\alpha$ -amylases. In the present disclosure, this family is termed  
30 "Fungamyl-like  $\alpha$ -amylase" and intends to indicate an  $\alpha$ -amylase which exhibits a high homology, i.e. more than 70%, such as 80% homologous (as defined herein) to the fungal  $\alpha$ -amylase derived from *Aspergillus oryzae*, commercially available as Fungamyl®, and the *A. niger*  $\alpha$ -amylase.

35 From the enclosed illustrations of the  $\alpha$ -amylase structure of a Termamyl-like  $\alpha$ -amylase and a comparison of said structure with the structure of a Fungamyl-like  $\alpha$ -amylase it is evident

that major differences exist between the two structures. In the method of the invention it is of particular interest to modify parts of the parent Termamyl-like  $\alpha$ -amylase, which belong to a region with large differences to the Fungamyl-like  $\alpha$ -amylase. In particular, it is of interest to modify the parent Termamyl-like  $\alpha$ -amylase in one or more of the following loops: loop 1, loop 2, loop 3 and/or loop 8 of the parent  $\alpha$ -amylase.

In the method of the third aspect of the invention it is of particular interest to modify loop 1, loop 2, loop 3 and/or loop 8 of the parent non-Termamyl-like  $\alpha$ -amylase to a closer resemblance to the similar loops of a Termamyl-like  $\alpha$ -amylase, such as Termamyl.

In the following specific types of variants are described which have been designed by use of the method of the invention.

#### Loop modifications

In order to change the substrate specificity of the parent  $\alpha$ -amylase to be modified it is relevant to consider loop modifications. For instance changing one or more of the loop structures of the Termamyl-like  $\alpha$ -amylase into a closer resemblance with the corresponding loop structure(s) of a non-Termamyl-like  $\alpha$ -amylase (such as a Fungamyl-like  $\alpha$ -amylase) it is contemplated that it is possible to change the substrate specificity in the direction of that of the non-Termamyl  $\alpha$ -amylase. In the following different types of loop modifications of interest are listed. It will be understood that the variants may have other changed properties in addition to the modified substrate specificity. It will be understood that the following modifications identified for a specific Termamyl-like  $\alpha$ -amylase are intended to include corresponding modifications in other equivalent positions of other Termamyl-like  $\alpha$ -amylases. Furthermore, it will be understood that, normally, the loop modification will comprise replacement of an entire loop structure or a substantial part thereof in, e.g., the Termamyl-

like  $\alpha$ -amylase, with the corresponding loop structure (or substantial part thereof) in a non-Termamyl-like  $\alpha$ -amylase.

*Loop2 modifications*

5 In one embodiment the invention relates to a variant of a parent Termamyl-like  $\alpha$ -amylase, in which variant at least one amino acid residue of the parent  $\alpha$ -amylase, which is/are present in a fragment corresponding to the amino acid fragment 44-57 of the amino acid sequence of SEQ ID No. 4, i.e. loop 2,  
10 has been deleted or replaced with one or more amino acid residues which is/are present in a fragment corresponding to the amino acid fragment 66-84 of the amino acid sequence shown in SEQ ID No. 10, or in which one or more additional amino acid residues has been added using the relevant part of SEQ ID No.  
15 10 or a corresponding part of another Fungamyl-like  $\alpha$ -amylase as a template.

The amino acid sequence shown in SEQ ID No. 10 is the amino acid sequence of the *A. oryzae*  $\alpha$ -amylase, i.e. a Fungamyl-like  
20  $\alpha$ -amylase. It will be understood that amino acid residues or fragments found in corresponding positions in other  $\alpha$ -amylases, in particular Fungamyl-like  $\alpha$ -amylases, may be used as a template for the construction of the variant according to the invention. The corresponding part in other homologous  $\alpha$ -  
25 amylases may easily be identified on the basis of a comparison of the amino acid sequences and/or three-dimensional structures of the respective  $\alpha$ -amylases.

For instance, the variant may be one, which, when the amino  
30 acid sequence of the variant is aligned most closely with the amino acid sequence of the said parent  $\alpha$ -amylase, occupies the same position as the portion from residue X to residue Y of SEQ ID No 4, the said region having at least 80% such as at least 90% sequence homology with the part of SEQ ID No 10 extending  
35 from residue Z to residue V of SEQ ID No 10, wherein X is the amino acid residue occupying position 44, 45, 46, 47 or 48 of SEQ ID No. 4,

Y is the amino acid residue occupying position 51, 52, 53, 54, 55, 56 or 57 of SEQ ID No. 4,  
Z is the amino acid residue occupying position 66, 67, 68, 69 or 70 of SEQ ID No. 10, and  
s V is the amino acid residue occupying position 78, 79, 80, 81, 82, 83 or 84 of SEQ ID No. 10.

In other words, the variant may be one in which an amino acid fragment X-Y of the parent  $\alpha$ -amylase, which corresponds to or  
10 is within the amino acid fragment 44-57 of SEQ ID No. 4, has been replaced with an amino acid fragment Z-V, which corresponds to or is within the amino acid fragment 66-84 of the amino acid sequence shown in SEQ ID No. 10, in X, Y, Z and V have the meaning indicated above.

15

A specific example of a variant according to this embodiment is a variant of a parent Termamyl-like  $\alpha$ -amylase, in which the amino acid fragment of the parent  $\alpha$ -amylase, which corresponds to amino acid residues 48-51 of SEQ ID No. 4, has been replaced  
20 with an amino acid fragment corresponding to amino acid residues 70-78 of the amino acid sequence shown in SEQ ID No. 10.

#### *Loop 3 modifications - limited alteration*

25 In another embodiment the invention relates to a variant of a parent Termamyl-like  $\alpha$ -amylase, in which variant at least one of the amino acid residues of the parent  $\alpha$ -amylase, which is/are present in an amino acid fragment corresponding to the amino acid fragment 195-202 of the amino acid sequence of SEQ  
30 ID No. 4, has been deleted or replaced with one or more of the amino acid residues which is/are present in an amino acid fragment corresponding to the amino acid fragment 165-177 of the amino acid sequence shown in SEQ ID No. 10, or in which one or more additional amino acid residues has been added using the  
35 relevant part of SEQ ID No. 10 or a corresponding part of another Fungamyl-like  $\alpha$ -amylase as a template.

For instance, the variant may be one in which an amino acid fragment X-Y of the parent  $\alpha$ -amylase which corresponds to or is within the amino acid fragment 195-202 of SEQ ID No. 4, has been replaced by an amino acid fragment Z-V, which corresponds to or is within the amino acid fragment 165-177 of the amino acid sequence shown in SEQ ID No. 10, in which

X is an amino acid residue corresponding to the amino acid occupying position 195 or 196 of SEQ ID No. 4,

10

Y is an amino acid residue corresponding to the amino acid occupying position 198, 199, 200, 201, or 202 of SEQ ID No. 4,

Z is an amino acid residue corresponding to the amino acid occupying position 165 or 166 of SEQ ID No. 10, and

V is an amino acid residue corresponding to the amino acid occupying position 173, 174, 175, 176 or 177 of SEQ ID No. 10.

20 Expressed in another manner, the variant according to this aspect may be one, which, when the amino acid sequence of variant is aligned most closely with the amino acid sequence of the said parent Termamyl-like  $\alpha$ -amylase, occupies the same position as the portion from residue X to residue Y of SEQ ID No 4, the said region having at least 80%, such as 90% sequence homology with the part of SEQ ID No 10 extending from residue Z to residue V of SEQ ID No 10, the meaning of X, Y, Z and V being as identified above.

30 A specific example of a variant according to this embodiment is a variant of a parent Termamyl-like  $\alpha$ -amylase, in which the amino acid fragment of the parent  $\alpha$ -amylase, which corresponds to amino acid residues 196-198 of SEQ ID No. 4, has been replaced with the amino acid fragment corresponding to amino acid residues 166-173 of the amino acid sequence shown in SEQ ID No. 10.

*Loop 3 modifications - complete domain B*

In a further embodiment the invention relates to a variant of a parent Termamyl-like  $\alpha$ -amylase, in which variant at least one of the amino acid residues of the parent  $\alpha$ -amylase, which is/are present in a fragment corresponding to the amino acid fragment 117-185 of the amino acid sequence of SEQ ID No. 4, has/have been deleted or replaced with one or more of the amino acid residues, which is/are present in an amino acid fragment corresponding to the amino acid fragment 98-210 of the amino acid sequence shown in SEQ ID No. 10, or in which one or more additional amino acid residues has been added using the relevant part of SEQ ID No. 10 or a corresponding part of another Fungamyl-like  $\alpha$ -amylase as a template.

For instance, the variant may be one, in which an amino acid fragment X-Y of the parent  $\alpha$ -amylase, which corresponds to or is within the amino acid fragment 117-185 of SEQ ID No. 4, has been replaced with an amino acid fragment Z-V, which corresponds to or is within the amino acid fragment 98-210 of the amino acid sequence shown in SEQ ID No. 10, in which variant

X is an amino acid residue corresponding to the amino acid occupying position 117, 118, 119, 120 or 121 of SEQ ID No. 4,

Y is an amino acid residue corresponding to the amino acid occupying position 181, 182, 183, 184 or 185 of SEQ ID No. 4,

Z is an amino acid residue corresponding to the amino acid occupying position 98, 99, 100, 101, 102 of SEQ ID No. 10, and

V is an amino acid residue corresponding to the amino acid occupying position 206, 207, 208, 209 or 210 of SEQ ID No. 10.

A specific example of a variant according to this embodiment is a variant of a parent  $\alpha$ -amylase, in which an amino acid fragment of the parent  $\alpha$ -amylase, which corresponds to amino acid residues 121-181 of SEQ ID No. 4, has been replaced with

the amino acid fragment corresponding to amino acid residues 102-206 of the amino acid sequence shown in SEQ ID No. 10.

In another embodiment the invention relates to a variant of a parent Termamyl-like  $\alpha$ -amylase, in which variant at least one of the amino acid residues of the parent  $\alpha$ -amylase, which is/are present in a fragment corresponding to the amino acid fragment 117-181 of the amino acid sequence of SEQ ID No. 4, has/have been deleted or replaced with one or more of the amino acid residues, which is/are present in an amino acid fragment corresponding to the amino acid fragment 98-206 of the amino acid sequence shown in SEQ ID No. 10, or in which one or more additional amino acid residues has been added using the relevant part of SEQ ID No. 10 or a corresponding part of another Fungamyl-like  $\alpha$ -amylase as a template.

For instance, the variant may be one, in which the amino acid fragment X-Y of the parent  $\alpha$ -amylase, which corresponds to or is within the amino acid fragment 117-177 of SEQ ID No. 4, has/have been replaced with an amino acid fragment Z-V, which corresponds to or is within the amino acid fragment 98-202 of the amino acid sequence shown in SEQ ID No. 10, in which variant

X is an amino acid residue corresponding to the amino acid occupying position 117, 118, 119, 120 or 121 of SEQ ID No. 4,

Y is an amino acid residue corresponding to the amino acid occupying position 174, 175, 176 or 177 of SEQ ID No. 4,

Z is an amino acid residue corresponding to the amino acid occupying position 98, 99, 100, 101, 102 of SEQ ID No. 10, and

V is an amino acid residue corresponding to the amino acid occupying position 199, 200, 201 or 202 of SEQ ID No. 10.

A specific example of a variant according to this embodiment of the invention is a variant, in which the amino acid fragment of

the parent  $\alpha$ -amylase, which corresponds to amino acid residues 121-174 of SEQ ID No. 4, has been replaced with the amino acid fragment corresponding to amino acid residues 102-199 of the amino acid sequence shown in SEQ ID No. 10.

5

*Loop 1 modifications - minimal addition*

In a further embodiment the present invention relates to a variant of a parent Termamyl-like  $\alpha$ -amylase, in which variant at least one of the amino acid residues of the parent  $\alpha$ -  
10 amylase, which is/are present in an amino acid fragment corresponding to the amino acid fragment 12-19 of the amino acid sequence of SEQ ID No. 4, has/have been deleted or replaced with one or more of the amino acid residues, which  
15 is/are present in an amino acid fragment which corresponds to the amino acid fragment 28-42 of SEQ ID No. 10, or in which one or more additional amino acid residues has/have been inserted using the relevant part of SEQ ID No. 10 or a corresponding part of another Fungamyl-like  $\alpha$ -amylase as a template.

20 For instance, the variant may be one, in which the amino acid fragment X-Y of the parent  $\alpha$ -amylase, which corresponds to or is within the amino acid fragment 12-19 of SEQ ID No. 4, has/have been replaced with an amino acid fragment Z-V, which corresponds to or is within the amino acid fragment 28-42 of  
25 the amino acid sequence shown in SEQ ID No. 10, in which variant

X is an amino acid residue corresponding to the amino acid occupying position 12, 13 or 14 of SEQ ID No. 4,

30

Y is an amino acid residue corresponding to the amino acid occupying position 15, 16, 17, 18 or 19 of SEQ ID No. 4,

Z is an amino acid residue corresponding to the amino acid  
35 occupying position 28, 29, 30, 31 or 32 of SEQ ID No. 10, and

V is an amino acid residue corresponding to the amino acid occupying position 38, 39, 40, 41 or 42 of SEQ ID No. 10.



A specific example of a variant according to this aspect of the invention is a variant, in which the amino acid fragment of the parent  $\alpha$ -amylase, which corresponds to amino acid residues 14-15 of SEQ ID No. 4, has been replaced with the amino acid s fragment corresponding to amino acid residues 32-38 of the amino acid sequence shown in SEQ ID No. 10.

*Loop 1 modifications - complete loop*

In a further embodiment the invention relates to a variant of 10 a parent Termamyl-like  $\alpha$ -amylase, in which variant at least one of the amino acid residues of the parent  $\alpha$ -amylase, which is present in a fragment corresponding to amino acid residues 7-23 of the amino acid sequence of SEQ ID No. 4, has/have been 15 deleted or replaced with one or more amino acid residues, which is/are present in an amino acid fragment corresponding to amino acid residues 13-45 of the amino acid sequence shown in SEQ ID No. 10, or in which one or more additional amino acid residues has/have been inserted using the relevant part of SEQ ID No. 10 or a corresponding part of another Fungamyl-like  $\alpha$ -amylase as 20 a template.

For instance, the variant may be one, in which the amino acid fragment X-Y of the parent  $\alpha$ -amylase, which corresponds to or is within the amino acid fragment 7-23 of SEQ ID No. 4, 25 has/have been replaced with an amino acid fragment Z-V, which corresponds to or is within the amino acid fragment 13-45 of the amino acid sequence shown in SEQ ID No. 10, in which variant

30 X is an amino acid residue corresponding to the amino acid occupying position 7 or 8 of SEQ ID No. 4,

Y is an amino acid residue corresponding to the amino acid occupying position 18, 19, 20, 21, 22 or 23 of SEQ ID No. 4, 35

Z is an amino acid residue corresponding to the amino acid occupying position 13 or 14 of SEQ ID No. 10, and

V is an amino acid residue corresponding to the amino acid occupying position 40, 41, 42, 43, 44 or 45 of SEQ ID No. 10.

A specific variant according to this embodiment is one, in which the amino acid fragment of the parent  $\alpha$ -amylase, which corresponds to amino acid residues 8-18 of SEQ ID No. 4, has been replaced with the amino acid fragment corresponding to amino acid residues 14-40 of the amino acid sequence shown in SEQ ID No. 10.

10

*Loop 8 modifications*

In a further embodiment the invention relates to a variant of a parent Termamyl-like  $\alpha$ -amylase, in which variant at least one of the amino acid residues of the parent  $\alpha$ -amylase, which is present in a fragment corresponding to amino acid residues 322-346 of the amino acid sequence of SEQ ID No. 2, has/have been deleted or replaced with one or more amino acid residues, which is/are present in an amino acid fragment corresponding to amino acid residues 291-313 of the amino acid sequence shown in SEQ ID No. 10, or in which one or more additional amino acid residues has/have been inserted using the relevant part of SEQ ID No. 10 or a corresponding part of another Fungamyl-like  $\alpha$ -amylase as a template.

For instance, the variant may be one, in which the amino acid fragment X-Y of the parent  $\alpha$ -amylase, which corresponds to or is within the amino acid fragment 322-346 of SEQ ID No. 2, has/have been replaced with an amino acid fragment Z-V, which corresponds to or is within the amino acid fragment 291-313 of the amino acid sequence shown in SEQ ID No. 10, in which variant

X is an amino acid residue corresponding to the amino acid occupying position 322, 323, 324 or 325 of SEQ ID No. 2,

35

Y is an amino acid residue corresponding to the amino acid occupying position 343, 344, 345 or 346 of SEQ ID No. 2,

Z is an amino acid residue corresponding to the amino acid occupying position 291, 292, 293 or 294 of SEQ ID No. 10, and

V is an amino acid residue corresponding to the amino acid occupying position 310, 311, 312 or 313 of SEQ ID No. 10.

A specific variant according to this aspect of the invention is one, in which the amino acid fragment of the parent  $\alpha$ -amylase, which corresponds to amino acid residues 325-345 of SEQ D No. 2, has been replaced with the amino acid fragment corresponding to amino acid residues 294-313 of the amino acid sequence shown in SEQ ID No. 10.

#### Ca<sup>2+</sup> dependency

It is highly desirable to be able to decrease the Ca<sup>2+</sup> dependency of a Termamyl-like  $\alpha$ -amylase. Accordingly, in a further aspect the invention relates to a variant of a parent Termamyl-like  $\alpha$ -amylase, which exhibits  $\alpha$ -amylase activity and which has a decreased Ca<sup>2+</sup> dependency as compared to the parent  $\alpha$ -amylase. The decreased Ca<sup>2+</sup> dependency has the functional result that the variant exhibits a satisfactory amylolytic activity in the presence of a lower concentration of calcium ion in the extraneous medium than is necessary for the parent enzyme and, for example, therefore is less sensitive than the parent to calcium ion-depleting conditions such as those obtained in media containing calcium-complexing agents (such as certain detergent builders).

The decreased Ca<sup>2+</sup> dependency of the variant of the invention may advantageously be achieved by increasing the Ca<sup>2+</sup> binding affinity of the parent Termamyl-like  $\alpha$ -amylase, in other words the stronger the Ca<sup>2+</sup> binding of the enzyme, the lower is the Ca<sup>2+</sup> dependency.

It is presently believed that amino acid residues located within 10Å from a sodium or calcium ion are involved in or are of importance for the Ca<sup>2+</sup> binding capability of the enzyme.

Accordingly, the variant according to this aspect of the invention is preferably one, which has been modified in one or more amino acid residues present within 10Å from a calcium and/or sodium ion identified in the three-dimensional Termamyl-like  $\alpha$ -amylase structure in such a manner that the affinity of the  $\alpha$ -amylase for calcium is increased.

The amino acid residues found within a distance of 10Å from the  $\text{Ca}^{2+}$  binding sites of the *B. licheniformis*  $\alpha$ -amylase with the amino acid sequence SEQ ID NO 2 were determined as described in Example 2 and are as follows:

V102, I103, N104, H105, K106, R125, W155, W157, Y158, H159, F160, D161, G162, T163, Y175, K176, F177, G178, K180, A181, W182, D183, W184, E185, V186, S187, N192, Y193, D194, Y195, L196, M197, Y198, A199, D200, I201, D202, Y203, D204, H205, P206, V208, A209, D231, A232, V233, K234, H235, I236, K237, F238, F240, L241, A294, A295, S296, T297, Q298, G299, G300, G301, Y302, D303, M304, R305, K306, L307, W342, F343, L346, Q393, Y394, Y396, H405, H406, D407, I408, V409, R413, E414, G415, D416, S417, V419, A420, N421, S422, G423, L424, I428, T429, D430, G431, P432, V440, G441, R442, Q443, N444, A445, G446, E447, T448, W449, I462, G475, Y480, V481, Q482, R483.

In order to construct a variant according to this aspect of the invention it is desirable to replace at least one of the above mentioned amino acid residues (or an amino acid residue occupying an equivalent position in another Termamyl-like  $\alpha$ -amylase than that defined by SEQ ID NO 2), which is contemplated to be involved in providing a non-optimal calcium binding, with any other amino acid residue which improves the  $\text{Ca}^{2+}$  binding affinity of the variant enzyme. In practice, the identification and subsequent modification of the amino acid residue is performed by the following method:

15

i) identifying an amino acid residue within 10Å from a  $\text{Ca}^{2+}$  binding site of a Termamyl-like  $\alpha$ -amylase structure, which from

structural or functional considerations is believed to be responsible for a non-optimal calcium ion interaction,

ii) constructing a variant in which said amino acid residue is replaced with another amino acid residue which from structural or functional considerations is believed to be important for establishing a higher  $\text{Ca}^{2+}$  binding affinity, and testing the  $\text{Ca}^{2+}$  dependency of the resulting Termamyl-like  $\alpha$ -amylase variant.

10 In the present context, the term "non-optimal calcium ion interaction" is intended to indicate that the amino acid residue in question is selected on the basis of a presumption that substituting said amino acid residue for another may improve a calcium ion binding interaction of the enzyme. For  
15 instance, the amino acid residue in question may be selected on the basis of one or more of the following considerations:

- to obtain an improved interaction between a calcium ion and an amino acid residue located near to the surface of the enzyme  
20 (as identified from the structure of the Termamyl-like  $\alpha$ -amylase). For instance, if the amino acid residue in question is exposed to a surrounding solvent, it may be advantageous to increase the shielding of said amino acid residue from the solvent so as to provide for a stronger interaction between  
25 said amino acid residue and a calcium ion. This can be achieved by replacing said residue (or an amino acid residue in the vicinity of said residue contributing to the shielding) by an amino acid residue which is more bulky or otherwise results in an improved shielding effect.

30  
- to stabilize a calcium binding site, for instance by stabilizing the structure of the Termamyl-like  $\alpha$ -amylase (e.g. by stabilizing the contacts between the A, B and C domains or stabilizing one or more of the domains as such). This may,  
35 e.g., be achieved by providing for a better coordination to amino acid side chains, which may, e.g., be obtained by replacing an N residue with a D residue and/or a Q residue with

an E residue (e.g. N104D), e.g. within 10Å, and preferably within 3 or 4Å, of a calcium binding site.

- to protect the calcium binding site or to improve the coordination between the calcium ion and the calcium binding site, e.g. by providing a stronger interaction between the ion and the binding site.

Before actually constructing a Termamyl-like  $\alpha$ -amylase variant according to the above principles it may be convenient to evaluate the contemplated amino acid modification by its accommodation into the Termamyl-like  $\alpha$ -amylase structure, e.g. into a model structure of the parent Termamyl-like  $\alpha$ -amylase.

Preferably, the amino acid residue to be modified is located within 8Å of a  $\text{Ca}^{2+}$  binding site residue, such as within 5Å of such residue. The amino acid residues within 8Å and 5Å, respectively, may easily be identified by an analogous method used for identifying amino acid residues within 10Å (cf. Example 2).

The following mutation is contemplated to be of particular interest with respect to decreasing the  $\text{Ca}^{2+}$  dependency of a Termamyl-like  $\alpha$ -amylase:

N104D (of the *B. licheniformis*  $\alpha$ -amylase SEQ ID NO 2, or an equivalent (N to D) mutation of an equivalent position in another Termamyl-like  $\alpha$ -amylase.)

In connection with substitutions of relevance for  $\text{Ca}^{2+}$  dependency, some other substitutions appear to be of importance in stabilizing the enzyme conformation (for instance the Domains A-B and/or Domains A-C interactions contributing to the overall stability of the enzyme) in that they may, e.g., enhance the strength of binding or retention of calcium ion or sodium ion at or within a calcium or sodium binding site, respectively, within the parent Termamyl-like  $\alpha$ -amylase.

It is desirable to stabilize the C-domain in order to increase the calcium stability and/or thermostability of the enzyme. In this connection the stabilization may result in a stabilization of the binding of calcium by the enzyme, and an improved  
5 contact between the C-domain and the A-domain (of importance for thermostability). The latter may be achieved by introduction of cystein bridges, salt bridges or increase hydrogen, hydrophobic and/or electrostatic interactions.

10 For instance, the C-domain of the *B. licheniformis*  $\alpha$ -amylase having the amino acid sequence shown in SEQ ID No. 2 may be stabilized by introduction of a cystein bridge between domain A and domain C, e.g. by introducing of the following mutations: A349C+I479C and/or L346C+I430C.

15

A salt bridge may be obtained by introduction of the following mutations:

N457D,E

N457D,E+K385R

20 F350D,E+I430R,K

F350D,E+I411R,K

The calcium site of Domain C may be stabilized by replacing the amino acid residues H408 and/or G303 with any other amino acid  
25 residue. Of particular interest is the following mutations:

H408Q,E,N,D and/or G303N,D,Q,E

which are contemplated to provide a better calcium binding or protection from calcium depletion.

30 Similar mutations may be introduced in equivalent positions of other Termamyl-like  $\alpha$ -amylases.

Other substitution mutations (relative to *B. licheniformis*  $\alpha$ -amylase, SEQ ID No. 2) which appear to be of importance,  
35 *inter alia*, in the context of reducing calcium dependency include the following: R23K, H156Y, A181T, A209V and G310D (or equivalent mutations in equivalent positions in another Termamyl-like  $\alpha$ -amylase). Substitutions of R214 and P345 with

other amino acids may also be of importancen in this connection.

Variants with altered activity at higher/lower pH

5

It is contemplated that it is possible to change the pH optima of a Termamyl-like  $\alpha$ -amylase or the enzymatic activity at a given pH by changing the pKa of the active site residues. This may be achieved, e.g. by changing the electrostatic interaction  
10 or hydrophobic interaction between functional groups of amino acid side chains of the amino acid residue to be modified and of its close surroundings. This may, e.g., be accomplished by the following method:

15 i) in a structure of the Termamyl-like  $\alpha$ -amylase in question to identifying an amino acid residue within 15Å from an active site residue, in particular 10Å from an active site residue, which amino acid residue is contemplated to be involved in electrostatic or hydrophobic interactions with an active site  
20 residue,

ii) replacing, in the structure, said amino acid residue with an amino acid residue which changes the electrostatic and/or hydrophobic surroundings of an active site residue and  
25 evaluating the accomodation of the amino acid residue in the structure,

iii) optionally repeating step i) and/or ii) until an amino acid replacement has been identified which is accomodated into  
30 the structure,

iv) constructing a Termamyl-like  $\alpha$ -amylase variant resulting from steps i), ii) and optionally iii) and testing the pH dependent enzymatic activity of interest of said variant.

35

In the above method it may be of particular relevance to add a positively charged residue within 5Å of a glutamate (thereby lowering the pKa of the glutamate from about 4.5 to 4), or to



add a negatively charged residue within 5 Å of a glutamate (thereby increasing the pKa to about 5), or to make similar modifications within a distance of about 5Å of a Histidine.

5 In a further aspect the invention relates to a variant of a Termamyl-like  $\alpha$ -amylase which exhibits a higher activity at a lower pH (e.g. compared to the pH optimum) than the parent  $\alpha$ -amylase. In particular, the variant comprises a mutation of an amino acid residue corresponding to at least one of the  
10 following positions of the *B. licheniformis*  $\alpha$ -amylase (SEQ ID NO 2):

E336, Q333, P331, I236, V102, A232, I103, L196

The following mutations are of particular interest:

15

E336R,K

Q333R,K

P331R,K

V102R,K,A,T,S,G;

20 I236K,R,N;

I103K,R;

L196K,R;

A232T,S,G;

25 or any combination of two or more of these variants or any combination of one or more of these variants with any of the other variants disclosed herein.

In a still further aspect the invention relates to a variant of  
30 a Termamyl-like  $\alpha$ -amylase which has a higher activity at a higher pH than the parent  $\alpha$ -amylase. In particular, the variant comprises a mutation of an amino acid residue corresponding to at least one of the following positions of the *B. licheniformis*  $\alpha$ -amylase (SEQ ID NO 2):

35

N236, H281, Y273

In particular, the variant comprises a mutation corresponding to at least one of the following mutations of the *B. licheniformis*  $\alpha$ -amylase (SEQ ID NO 2):

5 N326I,Y,F,L,V  
H281F,I,L  
Y273F,W

or any combination of two or more of these variants or any  
10 combination of one or more of these variants with any of  
the other variants disclosed herein.

A mutation which appears to be importance in relation to the  
specific activity of variants of the invention is a mutation  
15 corresponding to the substitution S187D in *B. licheniformis*  $\alpha$ -  
amylase (SEQ ID NO 2).

Variants with increased thermostability and/or altered  
temperature optimum

20

In a further desired aspect the invention relates to a variant  
of a parent Termamyl-like  $\alpha$ -amylase, which variant is the  
result of one or more amino acid residues having been deleted  
from, replaced or added to the parent  $\alpha$ -amylase so as to obtain  
25 an increased thermostability of the variant.

The Termamyl-like  $\alpha$ -amylase structure contains a number of  
unique internal holes, which may contain water, and a number of  
crevices. In order to increase the thermostability of the  $\alpha$ -  
30 amylase it may be desirable to reduce the number of holes and  
crevices (or reduce the size of the holes or crevices), e.g. by  
introducing one or more hydrophobic contacts, preferably  
achieved by introducing bulkier residues, in the vicinity or  
surroundings of the hole. For instance, the amino acid residues  
35 to be modified are those which are involved in the formation of  
the hole.

Accordingly, in a further aspect the present invention relates to a method of increasing the thermostability and/or altering the temperature optimum of a parent Termamyl-like  $\alpha$ -amylase, which method comprises

5

i) identifying an internal hole or a crevice of the parent Termamyl-like  $\alpha$ -amylase in the three-dimensional structure of said  $\alpha$ -amylase,

10 ii) replacing, in the structure, one or more amino acid residues in the neighbourhood of the hole or crevice identified in i) with another amino acid residue which from structural or functional considerations is believed to increase the hydrophobic interaction and to fill out or reduce the size of  
15 the hole or crevice,

iii) constructing a Termamyl-like  $\alpha$ -amylase variant resulting from step ii) and testing the thermostability and/or temperature optimum of the variant.

20 The structure used for identifying the hole or crevice of the parent Termamyl-like  $\alpha$ -amylase may be the structure identified in Appendix 1 or a model structure of the parent Termamyl-like  $\alpha$ -amylase built thereon.

25 It will be understood that the hole or crevice is identified by the amino acid residues surrounding the hole/crevice, and that modification of said amino acid residues are of importance for filling or reducing the size of the hole/crevice. The particular amino acid residues referred to below are those  
30 which in crystal structure have been found to flank the hole/crevice in question.

In order to fill (completely or partly) a major hole located between domain A and B, mutation to any other amino acid  
35 residue of an amino acid residue corresponding to one or more of the following residues of the *B. licheniformis*  $\alpha$ -amylase (SEQ ID NO 2) is contemplated:

L61, Y62, F67, K106, G145, I212, S151, R214, Y150, F143,  
R146

5

Of particular interest is a mutation to a more bulky amino acid residue than the amino acid residue of the parent enzyme.

Of particular interest is a variant of a Termamyl-like  $\alpha$ -  
10 amylase which comprises a mutation corresponding to the following mutations (using the numbering of *B. licheniformis*  $\alpha$ -amylase (SEQ ID NO 2):

L61W,V,F;

15 Y62W;

F67W;

K106R,F,W;

G145F,W

I212F,L,W,Y,R,K;

20 S151 replaced with any other amino acid residue and in particular with F,W,I or L;

R214W;

Y150R,K;

F143W; and/or

25 R146W.

In order to fill a hole in the vicinity of the active site mutation to any other amino acid residue of an amino acid residue corresponding to one or more of the following residues  
30 of the *B. licheniformis*  $\alpha$ -amylase (SEQ ID NO 2) is contemplated:

L241, I236.

35 Of interest is a mutation to a more bulky amino acid residue.

Of particular interest is a variant of a Termamyl-like  $\alpha$ -amylase which comprises a mutation corresponding to one or more of the following mutations in the *B. licheniformis*  $\alpha$ -amylase:

5 L241I,F,Y,W; and/or  
I236L,F,W,Y

In order to fill a hole in the vicinity of the active site mutation to any other amino acid residue of an amino acid  
10 residue corresponding to one or more of the following residues of the *B. licheniformis*  $\alpha$ -amylase (SEQ ID NO 2) is contemplated:

L7, V259, F284

15 Of interest is a mutation to a more bulky amino acid residue.

Of particular interest is a variant of a Termamyl-like  $\alpha$ -amylase which comprises a mutation corresponding to one or more of the following mutations in the *B. licheniformis*  $\alpha$ -amylase:

20

L7F,I,W  
V259F,I,L  
F284W

25 In order to fill a hole in the vicinity of the active site mutation to any other amino acid residue of an amino acid residue corresponding to one or more of the following residues of the *B. licheniformis*  $\alpha$ -amylase (SEQ ID NO 2) is contemplated:

30

F350, F343

Of interest is a mutation to a more bulky amino acid residue.

35 Of particular interest is a variant of a Termamyl-like  $\alpha$ -amylase which comprises a mutation corresponding to one or more of the following mutations in the *B. licheniformis*  $\alpha$ -amylase:  
F350W

F343W

In order to fill a hole in the vicinity of the active site mutation to any other amino acid residue of an amino acid  
5 residue corresponding to one or more of the following residues of the *B. licheniformis*  $\alpha$ -amylase (SEQ ID NO 2) is contemplated:

L427, V481

10

Of interest is a mutation to a more bulky amino acid residue.

Of particular interest is a variant of a Termamyl-like  $\alpha$ -amylase which comprises a mutation corresponding to one or more  
15 of the following mutations in the *B. licheniformis*  $\alpha$ -amylase:

L427F,L,W

V481,F,I,L,W

20 Variants with an altered cleavage pattern

In the starch liquefaction process it is desirable to use an  $\alpha$ -amylase which is capable of degrading the starch molecules into long branched oligo saccharides (like, e.g. the Fungamyl-like  $\alpha$ -amylases) rather than shorter branched oligo saccharides  
25 (like conventional Termamyl-like  $\alpha$ -amylases). The resulting very small branched oligosaccharides (panose precursors) cannot be hydrolyzed properly by pullulanases, which in the liquefaction process are used after the  $\alpha$ -amylases and before the amyloglucosidases. Thus, in the presence of panose  
30 precursors the action of amylo-glucoamylase ends up with a high degree of the small branched limiting-dextrin, the trisaccharide panose. The presence of panose lowers the saccharification yield significantly and is thus undesirable.

35 Thus, one aim of the present invention is to change the degradation characteristics of a Termamyl-like  $\alpha$ -amylase to that of a Fungamyl-like  $\alpha$ -amylases without at the same time reducing the thermostability of the Termamyl-like  $\alpha$ -amylase.

Accordingly, in a further aspect the invention relates to a variant of a Termamyl-like  $\alpha$ -amylase which has a reduced ability to cleave a substrate close to the branching point.

s The variant may suitably be constructed by a method which comprises

i) identifying the substrate binding area of the parent Termamyl-like  $\alpha$ -amylase in a model of the three-dimensional  
10 structure of said  $\alpha$ -amylase, (e.g. within a sphere of 4Å from the substrate binding site (as defined in the section above entitled "Substrate Binding Site"),

ii) replacing, in the model, one or more amino acid residues of  
15 the substrate binding area of the cleft identified in i), which is/are believed to be responsible for the cleavage pattern of the parent  $\alpha$ -amylase, with another amino acid residue which from structural considerations is believed to result in an altered substrate cleavage pattern, or deleting one or more  
20 amino acid residues of the substrate binding area contemplated to introduce favourable interactions to the substrate or adding one or more amino acid residues to the substrate binding area contemplated to introduce favourable interactions to the substrate, and

25 iii) constructing a Termamyl-like  $\alpha$ -amylase variant resulting from step ii) and testing the substrate cleavage pattern of the variant.

Of particular interest is a variant which cleaves an  
30 amylopectin substrate, from the reducing end, more than one glucose unit from the branching point, preferably more than two or three glucose units from the branching point, i.e. at a further distance from the branching point than that obtained by use of a wild type *B. licheniformis*  $\alpha$ -amylase.

35

Residues of particular interest in connection with this aspect of the invention correspond to the following residues of the *B. licheniformis*  $\alpha$ -amylase (SEQ ID NO 2): V54, D53, Y56, Q333,

G57, and the variants according to this aspect preferably comprises a mutation in one or more of these residues.

In particular, the variant comprises at least one of the following mutations, which are expected to prevent cleavage close to the branching point:

V54L, I, F, Y, W, R, K, H, E, Q

D53L, I, F, Y, W

10 Y56W

Q333W

G57all possible amino acid residues

A52amino acid residues larger than A, e.g. A52W, Y, L, F, I.

15 Variants of a fungal  $\alpha$ -amylase

In a still further embodiment the invention relates to a variant of a parent Fungamyl-like  $\alpha$ -amylase, in which variant at least one of the amino acid residues of the parent  $\alpha$ -  
20 amylase, which is/are present in an amino acid fragment corresponding to amino acid residues 291-313 of the amino acid sequence of SEQ ID No. 10, has/have been deleted or replaced with one or more of the amino acid residues, which is/are present in an amino acid fragment corresponding to amino acid  
25 residues 98-210 of the amino acid sequence shown in SEQ ID No. 4, or in which one or more additional amino acid residues has/have been inserted using the relevant part of SEQ ID No. 4 or a corresponding part of another Termamyl-like  $\alpha$ -amylase as a template.

30

For instance, the variant may be one, in which the amino acid fragment X-Y of the parent  $\alpha$ -amylase, which corresponds to or is within the amino acid fragment 117-185 of SEQ ID No. 10, has/have been replaced with an amino acid fragment Z-V, which  
35 corresponds to or is within the amino acid fragment 98-210 of the amino acid sequence shown in SEQ ID No. 4, in which variant



X is an amino acid residue corresponding to the amino acid occupying position 117, 118, 119, 120 or 121 of SEQ ID No. 10,

Y is an amino acid residue corresponding to the amino acid occupying position 181, 182, 183, 184 or 185 of SEQ ID No. 10,

Z is an amino acid residue corresponding to the amino acid occupying position 98, 99, 100, 101 or 102 of SEQ ID No. 4, and

10 V is an amino acid residue corresponding to the amino acid occupying position 206, 207, 208, 209 or 210 of SEQ ID No. 4.

A specific example of a variant according to this aspect of the invention is one, in which the amino acid fragment of the  
15 parent  $\alpha$ -amylase, which corresponds to amino acid residues 121-181 of SEQ ID No. 10, has been replaced with the amino acid fragment corresponding to amino acid residues 102-206 of the amino acid sequence shown in SEQ ID No. 4.

20 Another example of a variant according to this aspect of the invention is one, in which the amino acid fragment of the parent  $\alpha$ -amylase, which corresponds to amino acid residues 121-174 of SEQ ID No. 10, has been replaced with the amino acid fragment corresponding to amino acid residues 102-199 of the  
25 amino acid sequence shown in SEQ ID No. 4.

In a further embodiment the invention relates to a variant of a parent Fungamyl-like  $\alpha$ -amylase, in which an amino acid fragment corresponding to amino acid residues 181-184 of the  
30 amino acid sequence shown in SEQ ID No. 10 has been deleted.

#### General mutations in variants of the invention

It may be preferred that the variant of the invention or  
35 prepared in accordance with the method of the invention comprises one or more modifications in addition to those outlined above. Thus, it may be advantageous that one or more proline residues present in the part of the  $\alpha$ -amylase variant

having been modified is/are replaced with a non-proline residue which may be any of the possible, naturally occurring non-proline residues, and which preferably is an alanine, glycine, serine, threonine, valine or leucine.

5

Analogously, it may be preferred that one or more cysteine residues present in the amino acid residues with which the parent  $\alpha$ -amylase is modified are replaced with a non-cysteine residues such as serine, alanine, threonine, glycine, valine or  
10 leucine.

Furthermore, the variant of the invention may either as the only modification or in combination with any of the above outlined modifications be modified so that one or more Asp  
15 and/or Glu present in an amino acid fragment corresponding to the amino acid fragment 185-209 of SEQ ID No. 8 is replaced by an Asn and/or Gln, respectively. Also of interest is the modification of one or more of the Lys residues present in the Termamyl-like  $\alpha$ -amylase is replaced by an Arg present in an  
20 amino acid fragment corresponding to the amino acid fragment 185-209 of SEQ ID No. 8 is replaced by an Asn and/or Gln, respectively.

It will be understood that in accordance with the present  
25 invention variants may be prepared which carry two or more of the above outlined modifications. For instance, variants may be prepared which comprises a modification in the loop 1 and loop 2 region, a modification in loop 2 and limited loop 3, a modification in loop 1, loop 2, loop 3 and loop 8, etc.

30

Furthermore, it may be advantageous to introduce point-mutations in any of the variants described herein.

#### Methods of preparing $\alpha$ -amylase variants

35 Several methods for introducing mutations into genes are known in the art. After a brief discussion of the cloning of  $\alpha$ -amylase-encoding DNA sequences, methods for generating

mutations at specific sites within the  $\alpha$ -amylase-encoding sequence will be discussed.

Cloning a DNA sequence encoding an  $\alpha$ -amylase

5 The DNA sequence encoding a parent  $\alpha$ -amylase may be isolated from any cell or microorganism producing the  $\alpha$ -amylase in question, using various methods well known in the art. First, a genomic DNA and/or cDNA library should be constructed using chromosomal DNA or messenger RNA from the organism that pro-  
10 duces the  $\alpha$ -amylase to be studied. Then, if the amino acid sequence of the  $\alpha$ -amylase is known, homologous, labelled oligonucleotide probes may be synthesized and used to identify  $\alpha$ -amylase-encoding clones from a genomic library prepared from the organism in question. Alternatively, a labelled oligonu-  
15 cleotide probe containing sequences homologous to a known  $\alpha$ -amylase gene could be used as a probe to identify  $\alpha$ -amylase-encoding clones, using hybridization and washing conditions of lower stringency.

20 Yet another method for identifying  $\alpha$ -amylase-encoding clones would involve inserting fragments of genomic DNA into an expression vector, such as a plasmid, transforming  $\alpha$ -amylase-negative bacteria with the resulting genomic DNA library, and then plating the transformed bacteria onto agar containing a  
25 substrate for  $\alpha$ -amylase, thereby allowing clones expressing the  $\alpha$ -amylase to be identified.

Alternatively, the DNA sequence encoding the enzyme may be prepared synthetically by established standard methods, e.g.  
30 the phosphoroamidite method described by S.L. Beaucage and M.H. Caruthers (1981) or the method described by Matthes et al. (1984). In the phosphoroamidite method, oligonucleotides are synthesized, e.g. in an automatic DNA synthesizer, purified, annealed, ligated and cloned in appropriate vectors.

35

Finally, the DNA sequence may be of mixed genomic and synthetic origin, mixed synthetic and cDNA origin or mixed genomic and cDNA origin, prepared by ligating fragments of synthetic,

genomic or cDNA origin (as appropriate, the fragments corresponding to various parts of the entire DNA sequence), in accordance with standard techniques. The DNA sequence may also be prepared by polymerase chain reaction (PCR) using specific primers, for instance as described in US 4,683,202 or R.K. Saiki et al. (1988).

#### Site-directed mutagenesis

Once an  $\alpha$ -amylase-encoding DNA sequence has been isolated, and desirable sites for mutation identified, mutations may be introduced using synthetic oligonucleotides. These oligonucleotides contain nucleotide sequences flanking the desired mutation sites; mutant nucleotides are inserted during oligonucleotide synthesis. In a specific method, a single-stranded gap of DNA, bridging the  $\alpha$ -amylase-encoding sequence, is created in a vector carrying the  $\alpha$ -amylase gene. Then the synthetic nucleotide, bearing the desired mutation, is annealed to a homologous portion of the single-stranded DNA. The remaining gap is then filled in with DNA polymerase I (Klenow fragment) and the construct is ligated using T4 ligase. A specific example of this method is described in Morinaga et al. (1984). US 4,760,025 discloses the introduction of oligonucleotides encoding multiple mutations by performing minor alterations of the cassette. However, an even greater variety of mutations can be introduced at any one time by the Morinaga method, because a multitude of oligonucleotides, of various lengths, can be introduced.

Another method of introducing mutations into  $\alpha$ -amylase-encoding DNA sequences is described in Nelson and Long (1989). It involves the 3-step generation of a PCR fragment containing the desired mutation introduced by using a chemically synthesized DNA strand as one of the primers in the PCR reactions. From the PCR-generated fragment, a DNA fragment carrying the mutation may be isolated by cleavage with restriction endonucleases and reinserted into an expression plasmid.

#### Random mutagenesis

Random mutagenesis is suitably performed either as localized or region-specific random mutagenesis in at least three parts of the gene translating to the amino acid sequence shown in question, or within the whole gene.

5

For region-specific random mutagenesis with a view to improving the thermal stability of a parent Termamyl-like  $\alpha$ -amylase, codon positions corresponding to the following amino acid residues of the *B. licheniformis*  $\alpha$ -amylase (SEQ ID NO 2) may  
10 appropriately be targeted:

To improve the stability of the calcium site between Domain A and C

I428-A435

15 T297-L308

F403-V409

To improve the stability between domain A and B:

D180-D204

20 H156-T163

A232-F238

With a view to achieving improved binding of a substrate (i.e. improved binding of a carbohydrate species, such as amylose or  
25 amylopectin) by a Termamyl-like  $\alpha$ -amylase variant, modified (e.g. higher) substrate specificity and/or modified (e.g. higher) specificity with respect to cleavage (hydrolysis) of substrate, it appears that the following codon positions for the amino acid sequence shown in SEQ ID NO 2 (or equivalent  
30 codon positions for another parent Termamyl-like  $\alpha$ -amylase in the context of the invention) may particularly appropriately be targeted:

13-18

35 50-56

70-76

102-109

163-172

189-199

229-235

360-264

327-335

5

The random mutagenesis of a DNA sequence encoding a parent  $\alpha$ -amylase to be performed in accordance with step a) of the above-described method of the invention may conveniently be performed by use of any method known in the art.

10

For instance, the random mutagenesis may be performed by use of a suitable physical or chemical mutagenizing agent, by use of a suitable oligonucleotide, or by subjecting the DNA sequence to PCR generated mutagenesis. Furthermore, the random mutagenesis may be performed by use of any combination of these mutagenizing agents.

The mutagenizing agent may, e.g., be one which induces transitions, transversions, inversions, scrambling, deletions, and/or insertions.

Examples of a physical or chemical mutagenizing agent suitable for the present purpose include ultraviolet (UV) irradiation, hydroxylamine, N-methyl-N'-nitro-N-nitrosoguanidine (MNNG), O-methyl hydroxylamine, nitrous acid, ethyl methane sulphonate (EMS), sodium bisulphite, formic acid, and nucleotide analogues.

When such agents are used, the mutagenesis is typically performed by incubating the DNA sequence encoding the parent enzyme to be mutagenized in the presence of the mutagenizing agent of choice under suitable conditions for the mutagenesis to take place, and selecting for mutated DNA having the desired properties.

35

When the mutagenesis is performed by the use of an oligonucleotide, the oligonucleotide may be doped or spiked with the three non-parent nucleotides during the synthesis of the oligonucleo-

tide at the positions which are to be changed. The doping or spiking may be done so that codons for unwanted amino acids are avoided. The doped or spiked oligonucleotide can be incorporated into the DNA encoding the amylolytic enzyme by any published technique, using e.g. PCR, LCR or any DNA polymerase and ligase.

When PCR-generated mutagenesis is used, either a chemically treated or non-treated gene encoding a parent  $\alpha$ -amylase enzyme is subjected to PCR under conditions that increase the mis-incorporation of nucleotides (Deshler 1992; Leung et al., Technique, Vol.1, 1989, pp. 11-15).

A mutator strain of *E. coli* (Fowler et al., Molec. Gen. Genet., 193, 1974, pp. 179-191), *S. cerevisiae* or any other microbial organism may be used for the random mutagenesis of the DNA encoding the amylolytic enzyme by e.g. transforming a plasmid containing the parent enzyme into the mutator strain, growing the mutator strain with the plasmid and isolating the mutated plasmid from the mutator strain. The mutated plasmid may subsequently be transformed into the expression organism.

The DNA sequence to be mutagenized may conveniently be present in a genomic or cDNA library prepared from an organism expressing the parent amylolytic enzyme. Alternatively, the DNA sequence may be present on a suitable vector such as a plasmid or a bacteriophage, which as such may be incubated with or otherwise exposed to the mutagenizing agent. The DNA to be mutagenized may also be present in a host cell either by being integrated in the genome of said cell or by being present on a vector harboured in the cell. Finally, the DNA to be mutagenized may be in isolated form. It will be understood that the DNA sequence to be subjected to random mutagenesis is preferably a cDNA or a genomic DNA sequence.

In some cases it may be convenient to amplify the mutated DNA sequence prior to the expression step (b) or the screening step (c) being performed. Such amplification may be performed in

accordance with methods known in the art, the presently preferred method being PCR-generated amplification using oligonucleotide primers prepared on the basis of the DNA or amino acid sequence of the parent enzyme.

5

Subsequent to the incubation with or exposure to the mutagenizing agent, the mutated DNA is expressed by culturing a suitable host cell carrying the DNA sequence under conditions allowing expression to take place. The host cell used for this purpose may be one which has been transformed with the mutated DNA sequence, optionally present on a vector, or one which was carried the DNA sequence encoding the parent enzyme during the mutagenesis treatment. Examples of suitable host cells are the following: grampositive bacteria such as *Bacillus subtilis*,  
15 *Bacillus licheniformis*, *Bacillus lentus*, *Bacillus brevis*, *Bacillus stearothermophilus*, *Bacillus alkalophilus*, *Bacillus amyloliquefaciens*, *Bacillus coagulans*, *Bacillus circulans*, *Bacillus lautus*, *Bacillus megaterium*, *Bacillus thuringiensis*, *Streptomyces lividans* or *Streptomyces murinus*; and gramnegative  
20 bacteria such as *E.coli*.

The mutated DNA sequence may further comprise a DNA sequence encoding functions permitting expression of the mutated DNA sequence.

25

Localized random mutagenesis: the random mutagenesis may advantageously be localized to a part of the parent  $\alpha$ -amylase in question. This may, e.g., be advantageous when certain regions of the enzyme have been identified to be of particular importance for a given property of the enzyme, and when  
30 modified are expected to result in a variant having improved properties. Such regions may normally be identified when the tertiary structure of the parent enzyme has been elucidated and related to the function of the enzyme.

35

The localized random mutagenesis is conveniently performed by use of PCR-generated mutagenesis techniques as described above or any other suitable technique known in the art.



Alternatively, the DNA sequence encoding the part of the DNA sequence to be modified may be isolated, e.g. by being inserted into a suitable vector, and said part may subsequently be subjected to mutagenesis by use of any of the mutagenesis methods discussed above.

With respect to the screening step in the above-mentioned method of the invention, this may conveniently be performed by use of a filter assay based on the following principle:

10

A microorganism capable of expressing the mutated amylolytic enzyme of interest is incubated on a suitable medium and under suitable conditions for the enzyme to be secreted, the medium being provided with a double filter comprising a first protein-binding filter and on top of that a second filter exhibiting a low protein binding capability. The microorganism is located on the second filter. Subsequent to the incubation, the first filter comprising enzymes secreted from the microorganisms is separated from the second filter comprising the microorganisms. 15  
20 The first filter is subjected to screening for the desired enzymatic activity and the corresponding microbial colonies present on the second filter are identified.

The filter used for binding the enzymatic activity may be any protein binding filter e.g. nylon or nitrocellulose. The top-filter carrying the colonies of the expression organism may be any filter that has no or low affinity for binding proteins e.g. cellulose acetate or Durapore™. The filter may be pretreated with any of the conditions to be used for screening or may be treated during the detection of enzymatic activity. 25  
30

The enzymatic activity may be detected by a dye, fluorescence, precipitation, pH indicator, IR-absorbance or any other known technique for detection of enzymatic activity.

35

The detecting compound may be immobilized by any immobilizing agent e.g. agarose, agar, gelatine, polyacrylamide, starch, filter paper, cloth; or any combination of immobilizing agents.

$\alpha$ -Amylase activity is detected by Cibacron Red labelled amylopectin, which is immobilized on agarose. For screening for variants with increased thermal and high-pH stability, the filter with bound  $\alpha$ -amylase variants is incubated in a buffer  
5 at pH 10.5 and 60° or 65°C for a specified time, rinsed briefly in deionized water and placed on the amylopectin-agarose matrix for activity detection. Residual activity is seen as lysis of Cibacron Red by amylopectin degradation. The conditions are chosen to be such that activity due to the  $\alpha$ -amylase having the  
10 amino acid sequence shown in SEQ ID No.1 can barely be detected. Stabilized variants show, under the same conditions, increased colour intensity due to increased liberation of Cibacron Red.

15 For screening for variants with an activity optimum at a lower temperature and/or over a broader temperature range, the filter with bound variants is placed directly on the amylopectin-Cibacron Red substrate plate and incubated at the desired temperature (e.g. 4°C, 10°C or 30°C) for a specified time.  
20 After this time activity due to the  $\alpha$ -amylase having the amino acid sequence shown in SEQ ID No.1 can barely be detected, whereas variants with optimum activity at a lower temperature will show increase amylopectin lysis. Prior to incubation onto the amylopectin matrix, incubation in all kinds of desired  
25 media - e.g. solutions containing  $\text{Ca}^{2+}$ , detergents, EDTA or other relevant additives - can be carried out in order to screen for changed dependency or for reaction of the variants in question with such additives.

30

#### Testing of variants of the invention

The testing of variants of the invention may suitably be performed by determining the starch-degrading activity of the  
35 variant, for instance by growing host cells transformed with a DNA sequence encoding a variant on a starch-containing agarose plate and identifying starch-degrading host cells. Further testing as to altered properties (including specific activity,

substrate specificity, cleavage pattern, thermoactivation, pH optimum, pH dependency, temperature optimum, and any other parameter) may be performed in accordance with methods known in the art.

5

#### Expression of $\alpha$ -amylase variants

According to the invention, a DNA sequence encoding the variant produced by methods described above, or by any alternative methods known in the art, can be expressed, in enzyme form,  
10 using an expression vector which typically includes control sequences encoding a promoter, operator, ribosome binding site, translation initiation signal, and, optionally, a repressor gene or various activator genes.

15 The recombinant expression vector carrying the DNA sequence encoding an  $\alpha$ -amylase variant of the invention may be any vector which may conveniently be subjected to recombinant DNA procedures, and the choice of vector will often depend on the host cell into which it is to be introduced. Thus, the vector  
20 may be an autonomously replicating vector, i.e. a vector which exists as an extrachromosomal entity, the replication of which is independent of chromosomal replication, e.g. a plasmid, a bacteriophage or an extrachromosomal element, minichromosome or an artificial chromosome. Alternatively, the vector may be one  
25 which, when introduced into a host cell, is integrated into the host cell genome and replicated together with the chromosome(s) into which it has been integrated.

In the vector, the DNA sequence should be operably connected to  
30 a suitable promoter sequence. The promoter may be any DNA sequence which shows transcriptional activity in the host cell of choice and may be derived from genes encoding proteins either homologous or heterologous to the host cell. Examples of suitable promoters for directing the transcription of the DNA  
35 sequence encoding an  $\alpha$ -amylase variant of the invention, especially in a bacterial host, are the promoter of the lac operon of *E. coli*, the *Streptomyces coelicolor* agarase gene *dagA* promoters, the promoters of the *Bacillus licheniformis*  $\alpha$ -

amylase gene (*amyL*), the promoters of the *Bacillus stearothermophilus* maltogenic amylase gene (*amyM*), the promoters of the *Bacillus amyloliquefaciens*  $\alpha$ -amylase (*amyQ*), the promoters of the *Bacillus subtilis* *xylA* and *xylB* genes etc. For transcription in a fungal host, examples of useful promoters are those derived from the gene encoding *A. oryzae* TAKA amylase, *Rhizomucor miehei* aspartic proteinase, *A. niger* neutral  $\alpha$ -amylase, *A. niger* acid stable  $\alpha$ -amylase, *A. niger* glucoamylase, *Rhizomucor miehei* lipase, *A. oryzae* alkaline protease, *A. oryzae* triose phosphate isomerase or *A. nidulans* acetamidase.

The expression vector of the invention may also comprise a suitable transcription terminator and, in eukaryotes, polyadenylation sequences operably connected to the DNA sequence encoding the  $\alpha$ -amylase variant of the invention. Termination and polyadenylation sequences may suitably be derived from the same sources as the promoter.

The vector may further comprise a DNA sequence enabling the vector to replicate in the host cell in question. Examples of such sequences are the origins of replication of plasmids pUC19, pACYC177, pUB110, pE194, pAMB1 and pIJ702.

The vector may also comprise a selectable marker, e.g. a gene the product of which complements a defect in the host cell, such as the *dal* genes from *B. subtilis* or *B. licheniformis*, or one which confers antibiotic resistance such as ampicillin, kanamycin, chloramphenicol or tetracyclin resistance. Furthermore, the vector may comprise *Aspergillus* selection markers such as *amdS*, *argB*, *niaD* and *sC*, a marker giving rise to hygromycin resistance, or the selection may be accomplished by co-transformation, e.g. as described in WO 91/17243.

While intracellular expression may be advantageous in some respects, e.g. when using certain bacteria as host cells, it is generally preferred that the expression is extracellular. In general, the *Bacillus*  $\alpha$ -amylases mentioned herein comprise a

preregion permitting secretion of the expressed protease into the culture medium. If desirable, this preregion may be replaced by a different preregion or signal sequence, conveniently accomplished by substitution of the DNA sequences encoding the respective preregions.

The procedures used to ligate the DNA construct of the invention encoding an  $\alpha$ -amylase variant, the promoter, terminator and other elements, respectively, and to insert them into suitable vectors containing the information necessary for replication, are well known to persons skilled in the art (cf., for instance, Sambrook et al. (1989)).

The cell of the invention, either comprising a DNA construct or an expression vector of the invention as defined above, is advantageously used as a host cell in the recombinant production of an  $\alpha$ -amylase variant of the invention. The cell may be transformed with the DNA construct of the invention encoding the variant, conveniently by integrating the DNA construct (in one or more copies) in the host chromosome. This integration is generally considered to be an advantage as the DNA sequence is more likely to be stably maintained in the cell. Integration of the DNA constructs into the host chromosome may be performed according to conventional methods, e.g. by homologous or heterologous recombination. Alternatively, the cell may be transformed with an expression vector as described above in connection with the different types of host cells.

The cell of the invention may be a cell of a higher organism such as a mammal or an insect, but is preferably a microbial cell, e.g. a bacterial or a fungal (including yeast) cell.

Examples of suitable bacteria are grampositive bacteria such as *Bacillus subtilis*, *Bacillus licheniformis*, *Bacillus lentus*, *Bacillus brevis*, *Bacillus stearothermophilus*, *Bacillus alkalophilus*, *Bacillus amyloliquefaciens*, *Bacillus coagulans*, *Bacillus circulans*, *Bacillus lautus*, *Bacillus megaterium*, *Bacillus thuringiensis*, or *Streptomyces lividans* or *Streptomyces*

murinus, or gramnegative bacteria such as *E.coli*. The transformation of the bacteria may, for instance, be effected by protoplast transformation or by using competent cells in a manner known per se.

5

The yeast organism may favourably be selected from a species of *Saccharomyces* or *Schizosaccharomyces*, e.g. *Saccharomyces cerevisiae*. The filamentous fungus may advantageously belong to a species of *Aspergillus*, e.g. *Aspergillus oryzae* or *Aspergillus niger*. Fungal cells may be transformed by a process involving protoplast formation and transformation of the protoplasts followed by regeneration of the cell wall in a manner known per se. A suitable procedure for transformation of *Aspergillus* host cells is described in EP 238 023.

15

In a yet further aspect, the present invention relates to a method of producing an  $\alpha$ -amylase variant of the invention, which method comprises cultivating a host cell as described above under conditions conducive to the production of the variant and recovering the variant from the cells and/or culture medium.

The medium used to cultivate the cells may be any conventional medium suitable for growing the host cell in question and obtaining expression of the  $\alpha$ -amylase variant of the invention. Suitable media are available from commercial suppliers or may be prepared according to published recipes (e.g. as described in catalogues of the American Type Culture Collection).

The  $\alpha$ -amylase variant secreted from the host cells may conveniently be recovered from the culture medium by well-known procedures, including separating the cells from the medium by centrifugation or filtration, and precipitating proteinaceous components of the medium by means of a salt such as ammonium sulphate, followed by the use of chromatographic procedures such as ion exchange chromatography, affinity chromatography, or the like.

### Industrial Applications

The  $\alpha$ -amylase variants of this invention possesses valuable properties allowing for various industrial applications. In particular the enzyme variants finds potential applications as a component in washing, dishwashing and hard surface cleaning detergent compositions, but it may also be useful in the production of sweeteners and ethanol from starch and for textile desizing. Conditions for conventional starch converting processes and liquefaction and/or saccharification processes are described in for instance US Patent No. 3,912,590 and EP patent publications Nos. 252,730 and 63,909.

Production of sweetners from starch: A "traditional" process for conversion of starch to fructose syrups normally consists of three consecutive enzymatic processes, viz. a liquefaction process followed by a saccharification process and an isomerization process. During the liquefaction process, starch is degraded to dextrins by an  $\alpha$ -amylase (e.g. Termamyl<sup>®</sup>) at pH values between 5.5 and 6.2 and at temperatures of 95-160°C for a period of approx. 2h. In order to ensure an optimal enzyme stability under these conditions, 1mM of calcium is added (40 ppm free calcium ions).

After the liquefaction process the dextrins are converted into dextrose by addition of a glucoamylase (e.g. AMG<sup>™</sup>) and a debranching enzyme, such as an isoamylase or a pullulanase (e.g. Promozyme<sup>™</sup>). Before this step the pH is reduced to a value below 4.5, maintaining the high temperature (above 95°C), and the liquefying  $\alpha$ -amylase activity is denatured. The temperature is lowered to 60°C, and glucoamylase and debranching enzyme are added. The saccharification process proceeds for 24-72 hours.

After the saccharification process the pH is increased to a value in the range of 6-8, preferably pH 7.5, and the calcium is removed by ion exchange. The dextrose syrup is then converted into high fructose syrup using, e.g., an immobilized glucoseisomerase (such as Sweetzyme<sup>™</sup>).

At least 3 enzymatic improvements of this process could be obtained. All three improvements could be seen as individual benefits, but any combination (e.g. 1+2, 1+3, 2+3 or 1+2+3) could be employed:

5

Improvement 1. Reduction of the calcium dependency of the liquefying  $\alpha$ -amylase.

Addition of free calcium is required to ensure adequately high  
10 stability of the  $\alpha$ -amylase, but free calcium strongly inhibits the activity of the glucose isomerase and needs to be removed, by means of an expensive unit operation, to an extent which reduces the level of free calcium to below 3-5 ppm. Cost savings could be obtained if such an operation could be avoided  
15 and the liquefaction process could be performed without addition of free calcium ions.

To achieve that, a less calcium-dependent Termamyl-like  $\alpha$ -amylase which is stable and highly active at low  
20 concentrations of free calcium (< 40 ppm) is required. Such a Termamyl-like  $\alpha$ -amylase should have a pH optimum at a pH in the range of 4.5-6.5, preferably in the range of 4.5-5.5.

Improvement 2. Reduction of formation of unwanted Maillard  
25 products

The extent of formation of unwanted Maillard products during the liquefaction process is dependent on the pH. Low pH favours reduced formation of Maillard products. It would thus be  
30 desirable to be able to lower the process pH from around pH 6.0 to a value around pH 4.5; unfortunately, all commonly known, thermostable Termamyl-like  $\alpha$ -amylases are not very stable at low pH (i.e. pH < 6.0) and their specific activity is generally low.

35

Achievement of the above-mentioned goal requires a Termamyl-like  $\alpha$ -amylase which is stable at low pH in the range of



4.5-5.5 and at free calcium concentrations in the range of 0-40 ppm, and which maintains a high specific activity.

### Improvement 3.

5

It has been reported previously (US patent 5,234,823) that when saccharifying with *A. niger* glucoamylase and *B. acidopullulyticus* pullulanase, the presence of residual  $\alpha$ -amylase activity from the liquefaction process can lead to lower yields of dextrose if the  $\alpha$ -amylase is not inactivated before the saccharification stage. This inactivation can typically be carried out by adjusting the pH to below 4.3 at 95°C, before lowering the temperature to 60°C for saccharification.

15 The reason for this negative effect on dextrose yield is not fully understood, but it is assumed that the liquefying  $\alpha$ -amylase (for example Termamyl™ 120 L from *B. licheniformis*) generates "limit dextrans" (which are poor substrates for *B. acidopullulyticus* pullulanase) by hydrolysing 1,4-alpha-glucosidic linkages close to and on both sides of the branching points in amylopectin. Hydrolysis of these limit dextrans by glucoamylase leads to a build-up of the trisaccharide panose, which is only slowly hydrolysed by glucoamylase.

25 The development of a thermostable  $\alpha$ -amylase which does not suffer from this disadvantage would be a significant process improvement, as no separate inactivation step would be required.

30 If a Termamyl-like, low-pH-stable  $\alpha$ -amylase is developed, an alteration of the specificity could be an advantage needed in combination with increased stability at low pH.

The methodology and principles of the present invention make it possible to design and produce variants according to the invention having the required properties as outlined above.

### Detergent Compositions

According to the invention, the  $\alpha$ -amylase may typically be a component of a detergent composition. As such, it may be included in the detergent composition in the form of a non-dusting granulate, a stabilized liquid, or a protected enzyme.

5 Non-dusting granulates may be produced, e.g. as disclosed in US 4,106,991 and 4,661,452 (both to Novo Industri A/S) and may optionally be coated by methods known in the art. Examples of waxy coating materials are poly(ethylene oxide) products (polyethyleneglycol, PEG) with mean molar weights of 1000 to

10 20000, ethoxylated nonylphenols having from 16 to 50 ethylene oxide units; ethoxylated fatty alcohols in which the alcohol contains from 12 to 20 carbon atoms and in which there are 15 to 80 ethylene oxide units; fatty alcohols; fatty acids; and mono- and di- and triglycerides of fatty acids. Examples of

15 film-forming coating materials suitable for application by fluid bed techniques are given in patent GB 1483591. Liquid enzyme preparations may, for instance, be stabilized by adding a polyol such as propylene glycol, a sugar or sugar alcohol, lactic acid or boric acid according to established methods.

20 Other enzyme stabilizers are well known in the art. Protected enzymes may be prepared according to the method disclosed in EP 238,216.

The detergent composition of the invention may be in any convenient form, e.g. as powder, granules, paste or liquid. A

25 liquid detergent may be aqueous, typically containing up to 70% of water and 0-30% of organic solvent, or nonaqueous.

The detergent composition comprises one or more surfactants, each of which may be anionic, nonionic, cationic, or zwitter-

30 ionic. The detergent will usually contain 0-50% of anionic surfactant such as linear alkylbenzenesulfonate (LAS), alpha-olefinsulfonate (AOS), alkyl sulfate (fatty alcohol sulfate) (AS), alcohol ethoxysulfate (AEOS or AES), secondary alkane-sulfonates (SAS), alpha-sulfo fatty acid methyl esters, alkyl-

35 or alkenylsuccinic acid or soap. It may also contain 0-40% of nonionic surfactant such as alcohol ethoxylate (AEO or AE), carboxylated alcohol ethoxylates, nonylphenol ethoxylate, alkylpolyglycoside, alkyl dimethylamineoxide, ethoxylated fatty

acid monoethanolamide, fatty acid monoethanolamide, or polyhydroxy alkyl fatty acid amide (e.g. as described in WO 92/06154).

- 5 The detergent composition may additionally comprise one or more other enzymes, such as lipase, cutinase, protease, cellulase, peroxidase, e.g., laccase.

The detergent may contain 1-65% of a detergent builder or  
10 complexing agent such as zeolite, diphosphate, triphosphate, phosphonate, citrate, nitrilotriacetic acid (NTA), ethylenediaminetetraacetic acid (EDTA), diethylenetriaminepentaacetic acid (DTMPA), alkyl- or alkenylsuccinic acid, soluble silicates or layered silicates (e.g. SKS-6 from Hoechst). The  
15 detergent may also be unbuilt, i.e. essentially free of detergent builder.

The detergent may comprise one or more polymers. Examples are carboxymethylcellulose (CMC), poly(vinylpyrrolidone) (PVP),  
20 polyethyleneglycol (PEG), poly(vinyl alcohol) (PVA), polycarboxylates such as polyacrylates, maleic/acrylic acid copolymers and lauryl methacrylate/acrylic acid copolymers.

The detergent may contain a bleaching system which may comprise  
25 a  $H_2O_2$  source such as perborate or percarbonate which may be combined with a peracid-forming bleach activator such as tetraacetylenediamine (TAED) or nonanoxylbenzenesulfonate (NOBS). Alternatively, the bleaching system may comprise peroxy acids of e.g. the amide, imide, or sulfone  
30 type.

The enzymes of the detergent composition of the invention may be stabilized using conventional stabilizing agents, e.g. a polyol such as propylene glycol or glycerol, a sugar or sugar  
35 alcohol, lactic acid, boric acid, or a boric acid derivative as e.g. an aromatic borate ester, and the composition may be formulated as described in e.g. WO 92/19709 and WO 92/19708.

The detergent may also contain other conventional detergent ingredients such as e.g. fabric conditioners including clays, foam boosters, suds suppressors, anti-corrosion agents, soil-suspending agents, anti-soil redeposition agents, dyes, bactericides, optical brighteners, or perfume.

The pH (measured in aqueous solution at use concentration) will usually be neutral or alkaline, e.g. 7-11.

Particular forms of detergent compositions within the scope of the invention include:

1) A detergent composition formulated as a granulate having a bulk density of at least 600 g/l comprising

	Linear alkylbenzenesulfonate (calculated as acid)	7	-	12%
20	Alcohol ethoxysulfate (e.g. C <sub>12-18</sub> alcohol, 1-2 EO) or alkyl sulfate (e.g. C <sub>12-18</sub> )	1	-	4%
	Alcohol ethoxylate (e.g. C <sub>12-18</sub> alcohol, 7 EO)	5	-	9%
	Sodium carbonate (as Na <sub>2</sub> CO <sub>3</sub> )	14	-	20%
25	Soluble silicate (as Na <sub>2</sub> O, 2SiO <sub>2</sub> )	2	-	6%
	Zeolite (as NaAlSiO <sub>4</sub> )	15	-	22%
	Sodium sulfate (as Na <sub>2</sub> SO <sub>4</sub> )	0	-	6%
	Sodium citrate/citric acid (as C <sub>6</sub> H <sub>5</sub> Na <sub>3</sub> O <sub>7</sub> /C <sub>6</sub> H <sub>8</sub> O <sub>7</sub> )	0	-	15%
30	Sodium perborate (as NaBO <sub>3</sub> ·H <sub>2</sub> O)	11	-	18%
	TAED	2	-	6%
	Carboxymethylcellulose	0	-	2%
	Polymers (e.g. maleic/acrylic acid copolymer, PVP, PEG)	0	-	3%
35	Enzymes (calculated as pure enzyme protein)	0.0001	-	0.1%
	Minor ingredients (e.g. suds suppressors, perfume, optical brightener, photobleach)	0	-	5%

2) A detergent composition formulated as a granulate having a bulk density of at least 600 g/l comprising

5	Linear alkylbenzenesulfonate (calculated as acid)	6 - 11%
	Alcohol ethoxysulfate (e.g. C <sub>12-18</sub> alcohol, 1-2 EO or alkyl sulfate (e.g. C <sub>12-18</sub> ))	1 - 3%
10	Alcohol ethoxylate (e.g. C <sub>12-18</sub> alcohol, 7 EO)	5 - 9%
	Sodium carbonate (as Na <sub>2</sub> CO <sub>3</sub> )	15 - 21%
	Soluble silicate (as Na <sub>2</sub> O, 2SiO <sub>2</sub> )	1 - 4%
15	Zeolite (as NaAlSiO <sub>4</sub> )	24 - 34%
	Sodium sulfate (as Na <sub>2</sub> SO <sub>4</sub> )	4 - 10%
	Sodium citrate/citric acid (as C <sub>6</sub> H <sub>5</sub> Na <sub>3</sub> O <sub>7</sub> /C <sub>6</sub> H <sub>8</sub> O <sub>7</sub> )	0 - 15%
	Carboxymethylcellulose	0 - 2%
20	Polymers (e.g. maleic/acrylic acid copolymer, PVP, PEG)	1 - 6%
	Enzymes (calculated as pure enzyme protein)	0.0001 - 0.1%
25	Minor ingredients (e.g. suds suppressors, perfume)	0 - 5%

3) A detergent composition formulated as a granulate having a bulk density of at least 600 g/l comprising

30	Linear alkylbenzenesulfonate (calculated as acid)	5 - 9%
	Alcohol ethoxylate (e.g. C <sub>12-18</sub> alcohol, 7 EO)	7 - 14%
35	Soap as fatty acid (e.g. C <sub>12-18</sub> fatty acid)	1 - 3%
	Sodium carbonate (as Na <sub>2</sub> CO <sub>3</sub> )	10 - 17%
	Soluble silicate (as Na <sub>2</sub> O, 2SiO <sub>2</sub> )	3 - 9%
	Zeolite (as NaAlSiO <sub>4</sub> )	23 - 33%
	Sodium sulfate (as Na <sub>2</sub> SO <sub>4</sub> )	0 - 4%
40	Sodium perborate (as NaBO <sub>3</sub> ·H <sub>2</sub> O)	8 - 16%

	TAED	2	-	8%
	Phosphonate (e.g. EDTMPA)	0	-	1%
	Carboxymethylcellulose	0	-	2%
5	Polymers (e.g. maleic/acrylic acid copolymer, PVP, PEG)	0	-	3%
	Enzymes (calculated as pure enzyme protein)	0.0001	-	0.1%
10	Minor ingredients (e.g. suds suppressors, perfume, optical brightener)	0	-	5%

4) A detergent composition formulated as a granulate having a bulk density of at least 600 g/l comprising

15	Linear alkylbenzenesulfonate (calculated as acid)	8	-	12%
	Alcohol ethoxylate (e.g. C <sub>12-15</sub> alcohol, 7 EO)	10	-	25%
20	Sodium carbonate (as Na <sub>2</sub> CO <sub>3</sub> )	14	-	22%
	Soluble silicate (as Na <sub>2</sub> O, 2SiO <sub>2</sub> )	1	-	5%
	Zeolite (as NaAlSiO <sub>3</sub> )	25	-	35%
	Sodium sulfate (as Na <sub>2</sub> SO <sub>4</sub> )	0	-	10%
	Carboxymethylcellulose	0	-	2%
25	Polymers (e.g. maleic/acrylic acid copolymer, PVP, PEG)	1	-	3%
	Enzymes (calculated as pure enzyme protein)	0.0001	-	0.1%
30	Minor ingredients (e.g. suds suppressors, perfume)	0	-	5%

5) An aqueous liquid detergent composition comprising

	Linear alkylbenzenesulfonate (calculated as acid)	15	-	21%
35	Alcohol ethoxylate (e.g. C <sub>12-15</sub> alcohol, 7 EO or C <sub>12-15</sub> alcohol, 5 EO)	12	-	18%
	Soap as fatty acid (e.g. oleic acid)	3	-	13%
40	Alkenylsuccinic acid (C <sub>12-14</sub> )	0	-	13%

	Aminoethanol	8	- 18%
	Citric acid	2	- 8%
	Phosphonate	0	- 3%
	Polymers (e.g. PVP, PEG)	0	- 3%
5	Borate (as $B_2O_3$ )	0	- 2%
	Ethanol	0	- 3%
	Propylene glycol	8	- 14%
	Enzymes (calculated as pure enzyme protein)	0.0001	- 0.1%
10	Minor ingredients (e.g. dispersants, suds suppressors, perfume, optical brightener)	0	- 5%

6) An aqueous structured liquid detergent composition comprising

	Linear alkylbenzenesulfonate (calculated as acid)	15	- 21%
20	Alcohol ethoxylate (e.g. $C_{12-15}$ alcohol, 7 EO, or $C_{12-15}$ alcohol, 5 EO)	3	- 9%
	Soap as fatty acid (e.g. oleic acid)	3	- 10%
	Zeolite (as $NaAlSiO_3$ )	14	- 22%
	Potassium citrate	9	- 18%
25	Borate (as $B_2O_3$ )	0	- 2%
	Carboxymethylcellulose	0	- 2%
	Polymers (e.g. PEG, PVP)	0	- 3%
30	Anchoring polymers such as, e.g., lauryl methacrylate/acrylic acid copolymer; molar ratio 25:1; MW 3800	0	- 3%
	Glycerol	0	- 5%
	Enzymes (calculated as pure enzyme protein)	0.0001	- 0.1%
35	Minor ingredients (e.g. dispersants, suds suppressors, perfume, optical brighteners)	0	- 5%

7) A detergent composition formulated as a granulate having a bulk density of at least 600 g/l comprising

	Fatty alcohol sulfate	5	- 10%
5	Ethoxylated fatty acid monoethanol- amide	3	- 9%
	Soap as fatty acid	0	- 3%
	Sodium carbonate (as $\text{Na}_2\text{CO}_3$ )	5	- 10%
	Soluble silicate (as $\text{Na}_2\text{O} \cdot 2\text{SiO}_2$ )	1	- 4%
	Zeolite (as $\text{NaAlSiO}_4$ )	20	- 40%
10	Sodium sulfate (as $\text{Na}_2\text{SO}_4$ )	2	- 8%
	Sodium perborate (as $\text{NaBO}_3 \cdot \text{H}_2\text{O}$ )	12	- 18%
	TAED	2	- 7%
	Polymers (e.g. maleic/acrylic acid copolymer, PEG)	1	- 5%
15	Enzymes (calculated as pure enzyme protein)	0.0001	- 0.1%
	Minor ingredients (e.g. optical brightener, suds suppressors, per- fume)	0	- 5%

8) A detergent composition formulated as a granulate comprising

	Linear alkylbenzenesulfonate (calculated as acid)	8	- 14%
25	Ethoxylated fatty acid monoethanol- amide	5	- 11%
	Soap as fatty acid	0	- 3%
	Sodium carbonate (as $\text{Na}_2\text{CO}_3$ )	4	- 10%
	Soluble silicate (as $\text{Na}_2\text{O} \cdot 2\text{SiO}_2$ )	1	- 4%
	Zeolite (as $\text{NaAlSiO}_4$ )	30	- 50%
30	Sodium sulfate (as $\text{Na}_2\text{SO}_4$ )	3	- 11%
	Sodium citrate (as $\text{C}_6\text{H}_5\text{Na}_3\text{O}_7$ )	5	- 12%
	Polymers (e.g. PVP, maleic/acrylic acid copolymer, PEG)	1	- 5%
35	Enzymes (calculated as pure enzyme protein)	0.0001	- 0.1%
	Minor ingredients (e.g. suds suppressors, perfume)	0	- 5%



9) A detergent composition formulated as a granulate comprising

	Linear alkylbenzenesulfonate (calculated as acid)	6	- 12%
	Nonionic surfactant	1	- 4%
5	Soap as fatty acid	2	- 6%
	Sodium carbonate (as $\text{Na}_2\text{CO}_3$ )	14	- 22%
	Zeolite (as $\text{NaAlSiO}_4$ )	18	- 32%
	Sodium sulfate (as $\text{Na}_2\text{SO}_4$ )	5	- 20%
	Sodium citrate (as $\text{C}_6\text{H}_5\text{Na}_3\text{O}_7$ )	3	- 8%
10	Sodium perborate (as $\text{NaBO}_3 \cdot \text{H}_2\text{O}$ )	4	- 9%
	Bleach activator (e.g. NOBS or TAED)	1	- 5%
	Carboxymethylcellulose	0	- 2%
15	Polymers (e.g. polycarboxylate or PEG)	1	- 5%
	Enzymes (calculated as pure enzyme protein)	0.0001	- 0.1%
	Minor ingredients (e.g. optical brightener, perfume)	0	- 5%

10) An aqueous liquid detergent composition comprising

	Linear alkylbenzenesulfonate (calculated as acid)	15	- 23%
25	Alcohol ethoxysulfate (e.g. $\text{C}_{12-15}$ alcohol, 2-3 EO)	8	- 15%
	Alcohol ethoxylate (e.g. $\text{C}_{12-15}$ al- cohol, 7 EO, or $\text{C}_{12-15}$ alcohol, 5 EO)	3	- 9%
30	Soap as fatty acid (e.g. lauric acid)	0	- 3%
	Aminoethanol	1	- 5%
	Sodium citrate	5	- 10%
	Hydrotrope (e.g. sodium toluenesulfonate)	2	- 6%
35	Borate (as $\text{B}_2\text{O}_3$ )	0	- 2%
	Carboxymethylcellulose	0	- 1%
	Ethanol	1	- 3%
	Propylene glycol	2	- 5%

	Enzymes (calculated as pure enzyme protein)	0.0001 - 0.1%
5	Minor ingredients (e.g. polymers, dispersants, perfume, optical brighteners)	0 - 5%

11) An aqueous liquid detergent composition comprising

	Linear alkylbenzenesulfonate (calculated as acid)	20 - 32%
10	Alcohol ethoxylate (e.g. C <sub>12-15</sub> alcohol, 7 EO, or C <sub>12-15</sub> alcohol, 5 EO)	6 - 12%
	Aminoethanol	2 - 6%
	Citric acid	8 - 14%
15	Borate (as B <sub>2</sub> O <sub>3</sub> )	1 - 3%
20	Polymer (e.g. maleic/acrylic acid copolymer, anchoring polymer such as, e.g., lauryl methacrylate/acrylic acid copolymer)	0 - 3%
	Glycerol	3 - 8%
	Enzymes (calculated as pure enzyme protein)	0.0001 - 0.1%
25	Minor ingredients (e.g. hydro-tropes, dispersants, perfume, optical brighteners)	0 - 5%

12) A detergent composition formulated as a granulate having a bulk density of at least 600 g/l comprising

30	Anionic surfactant (linear alkylbenzenesulfonate, alkyl sulfate, alpha-olefinsulfonate, alpha-sulfo fatty acid methyl esters, alkanesulfonates, soap)	25 - 40%
35	Nonionic surfactant (e.g. alcohol ethoxylate)	1 - 10%
	Sodium carbonate (as Na <sub>2</sub> CO <sub>3</sub> )	8 - 25%
	Soluble silicates (as Na <sub>2</sub> O, 2SiO <sub>2</sub> )	5 - 15%
	Sodium sulfate (as Na <sub>2</sub> SO <sub>4</sub> )	0 - 5%
40	Zeolite (as NaAlSiO <sub>3</sub> )	15 - 28%
	Sodium perborate (as NaBO <sub>3</sub> ·4H <sub>2</sub> O)	0 - 20%

	Bleach activator (TAED or NOBS)	0 - 5%
	Enzymes (calculated as pure enzyme protein)	0.0001 - 0.1%
5	Minor ingredients (e.g. perfume, optical brighteners)	0 - 3%

13) Detergent formulations as described in 1) - 12) wherein all or part of the linear alkylbenzenesulfonate is replaced by (C<sub>12</sub>-C<sub>18</sub>) alkyl sulfate.

10

14) A detergent composition formulated as a granulate having a bulk density of at least 600 g/l comprising

	(C <sub>12</sub> -C <sub>18</sub> ) alkyl sulfate	9 - 15%
15	Alcohol ethoxylate	3 - 6%
	Polyhydroxy alkyl fatty acid amide	1 - 5%
	Zeolite (as NaAlSiO <sub>4</sub> )	10 - 20%
	Layered disilicate (e.g. SK56 from Hoechst)	10 - 20%
20	Sodium carbonate (as Na <sub>2</sub> CO <sub>3</sub> )	3 - 12%
	Soluble silicate (as Na <sub>2</sub> O, 2SiO <sub>2</sub> )	0 - 6%
	Sodium citrate	4 - 8%
	Sodium percarbonate	13 - 22%
	TAED	3 - 8%
25	Polymers (e.g. polycarboxylates and PVP*)	0 - 5%
	Enzymes (calculated as pure enzyme protein)	0.0001 - 0.1%
30	Minor ingredients (e.g. optical brightener, photo bleach, perfume, suds suppressors)	0 - 5%

15) A detergent composition formulated as a granulate having a bulk density of at least 600 g/l comprising

(C <sub>12</sub> -C <sub>18</sub> ) alkyl sulfate	4 - 8%
Alcohol ethoxylate	11 - 15%

	Soap	1	- 4%
	Zeolite MAP or zeolite A	35	- 45%
	Sodium carbonate (as $\text{Na}_2\text{CO}_3$ )	2	- 8%
	Soluble silicate (as $\text{Na}_2\text{O} \cdot 2\text{SiO}_2$ )	0	- 4%
5	Sodium percarbonate	13	- 22%
	TAED	1	- 8%
	Carboxymethyl cellulose	0	- 3%
	Polymers (e.g. polycarboxylates and PVP)	0	- 3%
10	Enzymes (calculated as pure enzyme protein)	0.0001	- 0.1%
	Minor ingredients (e.g. optical brightener, phosphonate, perfume)	0	- 3%

15 16) Detergent formulations as described in 1) - 15) which contain a stabilized or encapsulated peracid, either as an additional component or as a substitute for already specified bleach systems.

20 17) Detergent compositions as described in 1), 3), 7), 9) and 12) wherein perborate is replaced by percarbonate.

18) Detergent compositions as described in 1), 3), 7), 9), 12), 14) and 15) which additionally contain a manganese catalyst.  
 25 The manganese catalyst may, e.g., be one of the compounds described in "Efficient manganese catalysts for low-temperature bleaching", Nature 369, 1994, pp. 637-639.

19) Detergent composition formulated as a nonaqueous detergent  
 30 liquid comprising a liquid nonionic surfactant such as, e.g., linear alkoxylated primary alcohol, a builder system (e.g. phosphate), enzyme and alkali. The detergent may also comprise anionic surfactant and/or a bleach system.

35 The  $\alpha$ -amylase variant of the invention may be incorporated in concentrations conventionally employed in detergents. It is at present contemplated that, in the detergent composition of the invention, the  $\alpha$ -amylase may be added in an amount correspon-

ding to 0.00001-1 mg (calculated as pure enzyme protein) of  $\alpha$ -amylase per liter of wash liquor.

#### Dishwashing Composition

- 5 The dishwashing detergent composition comprises a surfactant which may be anionic, non-ionic, cationic, amphoteric or a mixture of these types. The detergent will contain 0-90% of non-ionic surfactant such as low- to non-foaming ethoxylated propoxylated straight-chain alcohols.
- 10 The detergent composition may contain detergent builder salts of inorganic and/or organic types. The detergent builders may be subdivided into phosphorus-containing and non-phosphorus-containing types. The detergent composition usually contains 1-90% of detergent builders.
- 15 Examples of phosphorus-containing inorganic alkaline detergent builders, when present, include the water-soluble salts especially alkali metal pyrophosphates, orthophosphates, and polyphosphates. An example of phosphorus-containing organic alkaline detergent builder, when present, includes the water-
- 20 soluble salts of phosphonates. Examples of non-phosphorus-containing inorganic builders, when present, include water-soluble alkali metal carbonates, borates and silicates as well as the various types of water-insoluble crystalline or amorphous aluminosilicates of which zeolites are the best-known
- 25 representatives.

Examples of suitable organic builders include the alkali metal, ammonium and substituted ammonium, citrates, succinates, malonates, fatty acid sulphonates, carboxymethoxy succinates,

30 ammonium polyacetates, carboxylates, polycarboxylates, aminopolycarboxylates, polyacetyl carboxylates and polyhydroxysulphonates.

Other suitable organic builders include the higher molecular

35 weight polymers and co-polymers known to have builder properties, for example appropriate polyacrylic acid, polymaleic acid and polyacrylic/polymaleic acid copolymers and their salts.

The dishwashing detergent composition may contain bleaching agents of the chlorine/bromine-type or the oxygen-type. Examples of inorganic chlorine/bromine-type bleaches are lithium, sodium or calcium hypochlorite and hypobromite as well as chlorinated trisodium phosphate. Examples of organic chlorine/bromine-type bleaches are heterocyclic N-bromo and N-chloro imides such as trichloroisocyanuric, tribromoisocyanuric, dibromoisocyanuric and dichloroisocyanuric acids, and salts thereof with water-solubilizing cations such as potassium and sodium. Hydantoin compounds are also suitable.

The oxygen bleaches are preferred, for example in the form of an inorganic persalt, preferably with a bleach precursor or as a peroxy acid compound. Typical examples of suitable peroxy bleach compounds are alkali metal perborates, both tetrahydrates and monohydrates, alkali metal percarbonates, persilicates and perphosphates. Preferred activator materials are TAED and glycerol triacetate.

The dishwashing detergent composition of the invention may be stabilized using conventional stabilizing agents for the enzyme(s), e.g. a polyol such as e.g. propylene glycol, a sugar or a sugar alcohol, lactic acid, boric acid, or a boric acid derivative, e.g. an aromatic borate ester.

The dishwashing detergent composition of the invention may also contain other conventional detergent ingredients, e.g. deflocculant material, filler material, foam depressors, anti-corrosion agents, soil-suspending agents, sequestering agents, anti-soil redeposition agents, dehydrating agents, dyes, bactericides, fluorescers, thickeners and perfumes.

Finally, the  $\alpha$ -amylase variant of the invention may be used in conventional dishwashing detergents, e.g. in any of the detergents described in any of the following patent publications:

EP 518719, EP 518720, EP 518721, EP 516553, EP 516554,

EP 516555, GB 2200132, DE 3741617, DE 3727911, DE 4212166,  
DE 4137470, DE 3833047, WO 93/17089, DE 4205071, WO 52/09680,  
WO 93/18129, WO 93/04153, WO 92/06157, WO 92/08777, EP 429124,  
WO 93/21299, US 5141664, EP 561452, EP 561446, GB 2234980,  
5 WO 93/03129, EP 481547, EP 530870, EP 533239, EP 554943,  
EP 346137, US 5112518, EP 318204, EP 318279, EP 271155,  
EP 271156, EP 346136, GB 2228945, CA 2006687, WO 93/25651,  
EP 530635, EP 414197, US 5240632.

## 10 EXAMPLES

### EXAMPLE 1

## 15 Example on Homology building of TERM

The overall homology of the *B. licheniformis*  $\alpha$ -amylase (in the following referred to as TERM) to other Termamyl-like  $\alpha$ -amylases is high and the percent similarity is extremely high.  
20 The similarity calculated for TERM to BSG (the *B. stearothermophilus*  $\alpha$ -amylase with SEQ ID NO 6), and BAN (the *B. amyloliquefaciens*  $\alpha$ -amylase with SEQ ID NO 4) using the University of Wisconsin Genetics Computer Group's program GCG gave 89% and 78%, respectively. TERM has a deletion of 2  
25 residues between residue G180 and K181 compared to BAN and BSG. BSG has a deletion of 3 residues between G371 and I372 in comparison with BAN and TERM. Further BSG has a C-terminal extension of more than 20 residues compared to BAN and TERM. BAN has 2 residues less and TERM has one residue less in the  
30 N-terminal compared to BSG.

The structure of the *B. licheniformis* (TERM) and of the *B. amyloliquefaciens*  $\alpha$ -amylase (BAN), respectively, was model built on the structure disclosed in Appendix 1 herein. The  
35 structure of other Termamyl-like  $\alpha$ -amylases (e.g. those disclosed herein) may be built analogously.

In comparison with the  $\alpha$ -amylase used for elucidating the present structure, TERM differs in that it lacks two residues around 178-182. In order to compensate for this in the model structure, the HOMOLGY program from BIOSYM was used to substitute the residues in equivalent positions in the structure (not only structurally conserved regions) except for the deletion point. A peptide bond was established between G179(G177) and K180(K180) in TERM(BAN). The close structural relationship between the solved structure and the model structure (and thus the validity of the latter) is indicated by the presence of only very few atoms found to be too close together in the model.

To this very rough structure of TERM was then added all waters (605) and ions (4 Calcium and 1 Sodium) from the solved structure (Appendix 1) at the same coordinates as for said solved structure using the INSIGHT program. This could be done with only few overlaps - in other words with a very nice fit. This model structure were then minimized using 200 steps of Steepest descent and 600 steps of Conjugated gradient (see Brooks et al 1983, J. Computational Chemistry 4, p.187-217). The minimized structure was then subjected to molecular dynamics, 5ps heating followed by up to 200ps equilibration but more than 35ps. The dynamics as run with the Verlet algorithm and the equilibration temperature 300K were kept using the Berendsen coupling to a waterbath (Berendsen et. al., 1984, J. Chemical Physics 81, p. 3684-3690). Rotations and translations were removed every picosecond. The potential energy became stable after appr. 35ps equilibration. A mean dynamics structure was extracted and can be used for further analysis.

## EXAMPLE 2

Determination of residues within 10Å from the ions present in the solved structure

The coordinates of Appendix 1 are read into the INSIGHT program provided by BIOSYM technologies. The spatial coordinates are



presented showing the bonds between the atoms. The ions are presented as well as the water atoms. The program package part of creating subset are used to create a 10Å subset around the Calcium and the Sodium ions in the structure using the command  
5 ZONE. All residues having an atom within the 10Å are compiled and written out by the LIST MOLECULE command. By giving the ions the name ium in the coordinate file a 10Å sphere around all atoms called ium is compiled. The specific residues identified in this manner are given further above in the  
10 section entitled "Ca<sup>2+</sup> dependency".

### EXAMPLE 3

#### Determination of cavities in the solved structure (Appendix 1)

15

The solved structure exhibits many internal holes and cavities. When analysing for such cavities the Connolly program is normally used (Lee, B. and Richards, F.M. (1971) J. Mol. Biol. 55, p. 379-400). The program uses a probe with radius to search  
20 the external and internal surface of the protein. The smallest hole observable in this way has the probe radius.

To analyse the solved structure a modified version of the Connolly program included in the program of INSIGHT were used.  
25 First the water molecules and the ions were removed by unmerging these atoms from the solved structure. By using the command MOLECULE SURFACE SOLVENT the solvent accessible surface area were calculated for all atoms and residues using a probe radius of 1.4Å, and displayed on the graphics screen together  
30 with the model of the solved structure. The internal cavities where then seen as dot surfaces with no connections to external surface.

Mutant suggestions for filling out the holes are given in the  
35 specification (in the section entitled "Variants with increased thermostability and/or altered temperature optimum"). By using the homology build structures or/and the sequence alignment

mutations for the homologous structures of TERM and BSG and BAN can be made.

#### EXAMPLE 4

5

Construction of Termamyl™ variants in accordance with the invention

Termamyl (SEQ ID NO. 2) is expressed in *B. subtilis* from a  
10 plasmid denoted pDN1528. This plasmid contains the complete gene encoding Termamyl, *amyL*, the expression of which is directed by its own promoter. Further, the plasmid contains the origin of replication, *ori*, from plasmid pUB110 and the *cat* gene from plasmid pC194 conferring resistance towards  
15 chloramphenicol. pDN1528 is shown in Fig. 9.

A specific mutagenesis vector containing a major part of the coding region of SEQ ID NO 1 was prepared. The important features of this vector, denoted pJeEN1, include an origin of  
20 replication derived from the pUC plasmids, the *cat* gene conferring resistance towards chloramphenicol, and a frameshift-containing version of the *bla* gene, the wild type of which normally confers resistance towards ampicillin (*amp<sup>r</sup>* phenotype). This mutated version results in an *amp<sup>s</sup>* phenotype.  
25 The plasmid pJeEN1 is shown in Fig. 10, and the *E. coli* origin of replication, *ori*, *bla*, *cat*, the 5'-truncated version of the Termamyl amylase gene, and selected restriction sites are indicated on the plasmid.

30 Mutations are introduced in *amyL* by the method described by Deng and Nickoloff (1992, Anal. Biochem. 200, pp. 81-86) except that plasmids with the "selection primer" (primer #6616; see below) incorporated are selected based on the *amp<sup>r</sup>* phenotype of transformed *E. coli* cells harboring a plasmid with a repaired  
35 *bla* gene, instead of employing the selection by restriction enzyme digestion outlined by Deng and Nickoloff. Chemicals and enzymes used for the mutagenesis were obtained from the

Chameleon™ mutagenesis kit from Stratagene (catalogue number 200509).

After verification of the DNA sequence in variant plasmids, the truncated gene, containing the desired alteration, is subcloned into pDN1528 as a *Pst*I-*Eco*RI fragment and transformed into a protease- and amylase-depleted *Bacillus subtilis* strain in order to express the variant enzyme.

- 10 The Termamyl variant V54W was constructed by the use of the following mutagenesis primer (written 5' to 3', left to right):

PG GTC GTA GGC ACC GTA GCC CCA ATC CGC TTG

- 15 The Termamyl variant A52W + V54W was constructed by the use of the following mutagenesis primer (written 5' to 3', left to right):

PG GTC GTA GGC ACC GTA GCC CCA ATC CCA TTG GCT CG

20

Primer #6616 (written 5' to 3', left to right; P denotes a 5' phosphate):

P CTG TGA CTG GTG AGT ACT CAA CCA AGT C

25

#### EXAMPLE 5

Saccharification in the presence of "residual"  $\alpha$ -amylase activity

30

Two appropriate Termamyl variants with altered specificity were evaluated by saccharifying a DE 10 (DE = dextrose equivalent) maltodextrin substrate with *A. niger* glucoamylase and *B. acidopullulyticus* pullulanase under conditions where the variant amylase was active.

Saccharification: Substrates for saccharification were prepared by dissolving 230 g DE 10 spray-dried maltodextrin, prepared

from common corn starch, in 450 ml boiling deionized water and adjusting the dry substance (DS) content to approximately 30% w/w. The pH was adjusted to 4.7 (measured at 60°C) and aliquots of substrate corresponding to 15 g dry weight were transferred to 50 ml blue cap glass flasks.

The flasks were then placed in a shaking water bath equilibrated at 60°C, and the enzymes added. The pH was readjusted to 4.7 where necessary.

10

The following enzymes were used:

Glucoamylase: AMG<sup>™</sup> (Novo Nordisk A/S); dosage 0.18 AG/g DS

Pullulanase: Promozyme<sup>™</sup> (Novo Nordisk A/S);

15 dosage 0.06 PUN/g DS

$\alpha$ -Amylases: Termamyl<sup>™</sup> (Novo Nordisk A/S); dosage 60 NU/g DS

Termamyl variant V54W; dosage 60 NU/g DS

Termamyl variant V54W + A52W; dosage 60 NU/g DS

20 2 ml samples were taken periodically. The pH of each sample was adjusted to about 3.0, and the sample was then heated in a boiling water bath for 15 minutes to inactivate the enzymes. After cooling, the samples were treated with approximately 0.1 g mixed-bed ion exchange resin (BIO-Rad 501-X (D)) for 30 minutes on a rotary mixer and then filtered. The carbohydrate composition of each sample was determined by HPLC. The following results were obtained after 72 hours [DP<sub>n</sub> denotes a dextrose (D-glucose) oligomer with n glucose units]:

30

$\alpha$ -amylase	%DP <sub>1</sub>	%DP <sub>2</sub>	%DP <sub>3</sub>	%DP <sub>4</sub>
None (control)	95.9	2.8	0.4	1.0
V54W	96.0	2.9	0.4	0.8
35 V54W + A52W	95.9	2.8	0.4	0.8
Termamyl <sup>™</sup>	95.6	2.8	0.8	0.8

It can be seen from the above results that compared with the control (no  $\alpha$ -amylase activity present during liquefaction), the presence of  $\alpha$ -amylase activity from variants V54W and V54W + A52W did not lead to elevated panose (DP<sub>3</sub>) levels. In contrast, Termamyl  $\alpha$ -amylase activity resulted in higher levels of panose and a subsequent loss of D-glucose (DP<sub>1</sub>) yield.

Thus, if  $\alpha$ -amylase variants V54W or V54W + A52W are used for starch liquefaction, it will not be necessary to inactivate the residual  $\alpha$ -amylase activity before the commencement of saccharification.

#### EXAMPLE 6

##### Calcium-binding affinity of $\alpha$ -amylase variants of the invention

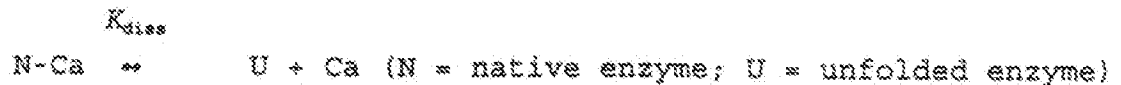
Unfolding of amylases by exposure to heat or to denaturants such as guanidine hydrochloride is accompanied by a decrease in fluorescence. Loss of calcium ions leads to unfolding, and the affinity of  $\alpha$ -amylases for calcium can be measured by fluorescence measurements before and after incubation of each  $\alpha$ -amylase (e.g. at a concentration of 10  $\mu$ g/ml) in a buffer (e.g. 50 mM HEPES, pH 7) with different concentrations of calcium (e.g. in the range of 1  $\mu$ M-100 mM) or of EGTA (e.g. in the range of 1-1000  $\mu$ M) [EGTA = 1,2-di(2-aminoethoxy)ethane-*N,N,N',N'*-tetraacetic acid] for a sufficiently long period of time (such as 22 hours at 55°C).

The measured fluorescence  $F$  is composed of contributions from the folded and unfolded forms of the enzyme. The following equation can be derived to describe the dependence of  $F$  on calcium concentration ( $[Ca]$ ):

$$F = \frac{[Ca]}{(K_{diss} + [Ca])} (\alpha_n - \beta_n \log([Ca])) + \frac{K_{diss}}{(K_{diss} + [Ca])} (\alpha_u - \beta_u \log([Ca]))$$

where  $\alpha_n$  is the fluorescence of the native (folded) form of the enzyme,  $\beta_n$  is the linear dependence of  $\alpha_n$  on the logarithm of

the calcium concentration (as observed experimentally),  $\alpha_0$  is the fluorescence of the unfolded form and  $\beta_0$  is the linear dependence of  $\alpha_0$  on the logarithm of the calcium concentration.  $K_{diss}$  is the apparent calcium-binding constant for an equilibrium process as follows:



In fact, unfolding proceeds extremely slowly and is irreversible. The rate of unfolding is dependent on calcium concentration, and the dependency for a given  $\alpha$ -amylase provides a measure of the Ca-binding affinity of the enzyme. By defining a standard set of reaction conditions (e.g. 22 hours at 55°C), a meaningful comparison of  $K_{diss}$  for different  $\alpha$ -amylases can be made. The calcium dissociation curves for  $\alpha$ -amylases in general can be fitted to the equation above, allowing determination of the corresponding values of  $K_{diss}$ .

The following values for  $K_{diss}$  were obtained for a parent Termamyl-like  $\alpha$ -amylase having the amino acid sequence shown in SEQ ID No. 1 of WO 95/26397 and for the indicated variant thereof according to the invention:

$\alpha$ -Amylase	$K_{diss}$ (mol/l)
L351C + M430C + T183* + G184*	$1.7 (\pm 0.5) \times 10^{-3}$
Parent	$3.5 (\pm 1.1) \times 10^{-3}$

It is apparent from the above that the calcium-binding affinity of the variant in question binds calcium significantly more strongly than the parent, and thereby has a correspondingly lower calcium dependency than the parent.

## REFERENCES CITED

- Klein, C., et al., *Biochemistry* 1992, 31, 8740-8746,
- s Mizuno, H., et al., *J. Mol. Biol.* (1993) 234, 1282-1283,
- Chang, C., et al., *J. Mol. Biol.* (1993) 229, 235-238,
- Larson, S.B., *J. Mol. Biol.* (1994) 235, 1560-1584,
- 10 Lawson, C.L., *J. Mol. Biol.* (1994) 236, 590-600,
- Qian, M., et al., *J. Mol. Biol.* (1993) 231, 785-799,
- 15 Brady, R.L., et al., *Acta Crystallogr. sect. B*, 47, 527-535,
- Swift, H.J., et al., *Acta Crystallogr. sect. B*, 47, 535-544
- A. Kadziola, Ph.D. Thesis: "An alpha-amylase from Barley and  
20 its Complex with a Substrate Analogue Inhibitor Studied by X-  
ray Crystallography", Department of Chemistry University of  
Copenhagen 1993
- MacGregor, E.A., *Food Hydrocolloids*, 1987, Vol.1, No. 5-6, p.
- 25 B. Diderichsen and L. Christiansen, Cloning of a maltogenic  $\alpha$ -  
amylase from *Bacillus stearothermophilus*, *FEMS Microbiol. let-*  
*ters* 56: pp. 53-60 (1988)
- 30 Hudson et al., *Practical Immunology*, Third edition (1989),  
Blackwell Scientific Publications,
- Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd  
Ed., Cold Spring Harbor, 1989
- 35 S.L. Beaucage and M.H. Caruthers, Tetrahedron Letters 22, 1981,  
pp. 1859-1869

Matthes et al., The EMBO J. 3, 1984, pp. 801-805.

R.K. Saiki et al., Science 239, 1988, pp. 487-491.

s Morinaga et al., (1984, Biotechnology 2:646-639)

Nelson and Long, Analytical Biochemistry 180, 1989, pp. 147-151

Hunkapiller et al., 1984, Nature 310:105-111

10

R. Higuchi, B. Krummel, and R.K. Saiki (1988). A general method of *in vitro* preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions. Nucl. Acids Res. 16:7351-7367.

15

Dubnau et al., 1971, J. Mol. Biol. 56, pp. 209-221.

Gryczan et al., 1978, J. Bacteriol. 134, pp. 318-329.

S.D. Erlich, 1977, Proc. Natl. Acad. Sci. 74, pp. 1680-1682.

20

Boel et al., 1990, Biochemistry 29, pp. 6244-6249.



## SEQUENCE LISTING

In the following SEQ ID Nos. 1, 3, 5 the 5', coding sequence and 3' sequence of the relevant  $\alpha$ -amylase genes are illustrated. The 5' sequence is the first separate part of the sequence written with lower case letters, the coding sequence is the intermediate part of the sequence, where the signal sequence is written with lower case letters and the sequence encoding the mature  $\alpha$ -amylase is written with upper case letters, and the 3' sequence is the third separate part of the sequence written with lower case letters.

## SEQ ID No. 1

15

cggaagattggaagtacaaaaataagcaaaagattgtcaatcatgtcatgagccatgcgg-  
gagacggaaaaatcgtctta atgcacgatattttatgcaacgttcgcagatgctgctgaa-  
gagattattaaaaagctgaaagcaaaaggctatcaattggc aactgtatctcagcttga-  
agaagtgaagaagcagagaggctattgaataaatgagtagaagcgccatatacggcgcttttc  
20 ttttggagaataatagggaaaatggtacttgttaaaaattcggaatatattatacaacatc-  
atatgtttcacattgaaa ggggaggagaatc

atgaaacaacaaaaacggctttacgcccagattgctgacgctgttattttgcgctcatcttctt-  
gctgc ctcatctetgcagcagcggcgGCAAATCTTAATGGGACGCTGATGCAGTATTTT-  
25 GAATGGTACATGCCCAATGACGGCCAA CATTGGAGGCGTTTGCAAACGACTCGGCATAT-  
TTGGCTGAACACGGTATTACTGCCGTCTGGATTCCCCCGGCATATAA GGGAACGAGC-  
CAAGCGGATGTGGGCTACGGTGCTTACGACCTTTATGATTTAGGGGAGTTTCATCAAAAAG-  
GGACGGTTC GGACAAAGTACGGCACAAAAGGAGAGCTGCAATCTGCGATCAAAAGTCTTC-  
ATTCCCGCGACATTAACGTTTACGGGGAT GTGGTCATCAACCACAAAGGCGGCGCTGA-  
30 TGCGACCGAAGATGTAACCGCGGTTGAAGTCGATCCCGCTGACCGCAACCG CGTAATTT-  
CAGGAGAACACCTAATTAAAGCCTGGACACATTTTTCATTTTCCGGGGCGCGGCAGCACATA-  
CAGCGATTTTA AATGGCATTGGTACCATTTTGACGGAACCGATTGGGACGAGTCCCGAAA-  
GCTGAACCGCATCTATAAGTTTCAAGGAAAG GCTTGGGATTGGGAAGTTTCCAATGAA-  
AACGGCAACTATGATTATTTGATGTATGCCGACATCGATTATGACCATCCTGA TGTCGCAG-  
35 CAGAAATTAAGAGATGGGGCACTTGGTATGCCAATGAACTGCAATTGGACGGTTTCCGCTTT-  
GATGCTGTCA AACACATTAAATTTTCTTTTTTGGGGATTGGGTTAATCATGTGAGGGA-  
AAAAACGGGGGAAGGAAATGTTTACGGTAGCT GAATATTGGCAGAAAT-  
GACTTGGGCGCGCTGGAAAATTTGAACAAAACAAATTTTAATCATTTCAGTGTTCAC-

GTGCC GCTTCATTATCAGTTCCATGCTGCATCGACACAGGGAGGCGGCTATGATATGAG-  
 GAAATTGCTGAACGGTACGGTCGTTT CCAAGCATCCGTTGAAATCGGTTACATTTGTGCG-  
 ATAACCATGATACACAGCCGGGGCAATCGCTTGAGTCGACTGTCCAA ACATGGTTTAAG-  
 CCGCTTGCTTACGCTTTTATTCTCACAAGGGAATCTGGATACCCTCAGGTTTTCTACGGG-  
 5 GATATGTA CGGGACGAAAGGAGACTCCAGCGCGAAATTCTGCTTGAAACACAAAAT-  
 TGAACCGATCTTAAAAGCGAGAAAACAGT ATGCGTACGGAGCACAGCATGATTATTTGAC-  
 CACCATGACATTGTCGGCTGGACAAGGGAAGGCGACAGCTCGGTTGCA AATTCAGGTTTGG-  
 CGGCATTAATAACAGACGGACCCGGTGGGGCAAAGCGAATGTATGTCGGCCCGGCA-  
 AAACGCCCGGTGA GACATGGCATGACATTACCGGAAACCGTTCGGAGCCGGTTGTCATCA-  
 10 ATTCGGAAGGCTGGGGAGAGTTTCACGTAAACG GCGGGTCCGTTTCAATTTATGTTCAA-  
 GATAG

aagagcagagaggacggattttcctgaaggaaatccgtttttttatattt

15 SEQ ID No. 2

ANLNGTLMQYFEWYMPNDGQHWRLQND SAYLAEHGITAV  
 WIPPAYKGT SQADVGYGAYDLYDLGEFHQKGT VRTKYGTK  
 GELQSAIKSLHSRDINVYGDV VINHKGGADATEDVTAVEV  
 20 DPADNRNVISGEHLIKAWTHFHFPGRGSTYSDFKWHWYHF  
 DGTDWDESRLNRIYKFQKAWDWEVSNENGN YDYL MYAD  
 IDYDHPDVAAEIKRWGTWYANELQLDGFRLDAVKHIKFSF  
 LRDWVNHVREKTGKEMFTVAEYWQNDLGALENYLNKTNFN  
 HSVFDVPLHYQFHAAS TQGGGYDMRKLLNGTVVSKHPLKS  
 25 VTFVDNHD TQPGQSLESTVQTWFKPLAYAFILTRESGYPO  
 VFYGD MYGT KGDSQREIPALKHKIEPILKARKQYAYGAQH  
 DYFDHHDIVGWTR EGDSSVANSGLAALITDGP GGAKRMVY  
 GRQNA GETWHDITGNRSEPVVINSEGWGEFHVNGGSVSIY  
 VQR

30

SEQ ID No. 3

gccccgcacatacgaaaaagactggctgaaaacattgagcctttgatgactgatgatttgg-  
 35 ctgaagaagtggatcgattg tttgagaaaaagaagaccataaaaaataccttgtctgt-  
 catcagacaggggtattttttatgctgtccagactgtccgct gtgtaaaaataaggaata-  
 aaggggggttgttattttttactgatatgtaaaaatataatttgtataagaaaatcagagggg  
 agagggaaac